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

NISTEP Seminar, Tokyo, Japan

20:50 - 21:20pm ET on December 13, 2022
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

Places & Spaces: Mapping Science Exhibit

1st Decade (2005-2014)

Maps

2nd Decade (2015-2024)

Macrosopes

Maps

- Iteration I (2005)
- Iteration III (2007)
- Iteration V (2009)
- Iteration VII (2011)
- Iteration IX (2013)

Macrosopes

- Iteration IV (2008)
- Iteration VI (2010)
- Iteration VIII (2012)
- Iteration X (2014)

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

http://scimaps.org
A Topic Map of NIH Grants 2007

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

The National Institutes of Health (NIH) is organized as a network of Institutes and Centers whose missions are primarily focused on distinct diseases. However, disease etiologies and therapies float 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 a topic map concept for navigating this landscape, online at wwwnihmaps.org. Institute abbreviations can be found at nih.gov/od/foa.

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

A core of this map focused on cardiac vascular functions and dysfunctions. Cardiac failure (primarily funded by NHLBI) is topologically clustered next to smoking (NIEHS), since these are often co-factor molecules in a systemic metabolic environment associated with ischemia, which results from restricted blood supply. Also localized in this area are grants focused on heart failure (NHLBI), signaling a unique mechanism pathway for vasodilation, and grants on hemodynamics, indoor cell disease, and nephrons.

Neural Circuits Research

An area of this map focused on neural circuits, which shows the identity of topics and NIH Institutes that fund research in this area, such as: Cardiac Temporary Respirations, predominantly funded by NHLBI, visual processing, primarily funded by NICHD and NIH, primarily funded by NINDS. For color coding, see label in the upper-left text.

National Cancer Institute (NCI)

- Oncology: Clinical Trials
- Cancer Treatment
- Cancer Therapy
- Cancer Genomics
- Risk Factor Analysis
- Cancer Chemotherapy
- Metabolomics
- Epidemiology
- Prediction: Prognosis
- Cancer Chemoprevention

National Institute of General Medical Sciences (NIGMS)

- Molecular Sequencing
- X-ray Crystallography
- Protein Structure
- Computational Models
- Yeast Biology
- Metabolomics
- Enzymatic Mechanisms
- Protein Complexes
- Invertebrate/Vertebrate Genetics
- Cell Division

National Heart, Lung, and Blood Institute (NHLBI)

- Cardiac Failure
- Pulmonary Injury
- Genetic Linkage Analysis
- Cardiovascular Disease
- Atherosclerosis
- Hematopoiesis
- Blood Pressure
- Asthma/Allergic Asthma Disease
- Gene Association
- Lipidomics

National Institute of Mental Health (NIMH)

- Mental Disorders
- Schizophrenia
- Behavioral Intervention Studies
- Mental Health
- Depression
- Cognitive-Behavioral Therapy
- AIDS Prevention
- Genetic Linkage Analysis
- Alzheimer
- Childhood

The Structure of Science

I.10 The Structure of Science - Kevin W. Boyack and Richard Klavans - 2005

The Medica Sciences include broad therapeutic studies and targeted areas of research (e.g. centre directed systems, oncology, pharmacogenomics, etc.). Unlike Physics and Chemistry, the medical disciplines are more specialized, suggesting a more multidisciplinary approach to research. The transition into Life Sciences (see next page) Science and Biochemistry is gradual.
Impact

The United States Patent and Trademark Office classifies and indexes patents using a hierarchical category structure. In this section, we discuss a taxonomy of patent data that groups related patents together into categories and subcategories.

III. Science-Related Wikipedian Activity

Bruce W. Herr II, Todd M. Holloway, Elisha F. Hardy, Katy Börner, and Kevin Boyack - 2007
Check out our **Zoom Maps** online!

Visit [scimaps.org](http://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/macroscope-kiosk
Smelly Maps – Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello – 2015
This is the Roanoke (Raleigh) megaregion.

Opportunity Atlas
Different zip codes, different outcomes

Mapping Inequality
Unequal by design

Atlas of Surveillance
One nation, under observation

Virus Explorer
Bugs in the system
Acknowledgements

Exhibit Curators

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

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

http://scimaps.org

Exhibit Advisory Board

Gary Berg-Cross
Cognitive psychologist (PhD, SUNY-Stony Brook), Potomac, MD, USA

André Skupin
Associate Professor of Geography at San Diego State University, California

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 DeVraco
Media X 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 Leibniz 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)
Call for Macroscopes: 19th 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
https://mitpress.mit.edu/books/atlas-forecasts
I am deeply grateful to all those who helped to make possible this atlas and the exhibit maps it features. Financial support came from the National Science Foundation under Grants No. DRL-1228608, OCI-0908204, SBE-0718911, and CBET-0831836; the National Institutes of Health under Grants No. U01 GM099959, R31-D034254, and U01-RR029822; the Jones S. MacDonald 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 CWTS, Andrea Schaarboek and Jeannette Haagsma of Maastricht, and Peter Doreen, Linda Reijnders, and Lucas Pauwerts of DANS for their support.

Part 2, “Envisioning Science,” benefited deeply from any teaching of relevant courses at Indiana University over the last 14 years, including teaching Information Visualizations 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 who helped to organize and run the exhibit. Exhibit advisors for the maps featured in this book include: Deborah MacPherson (Aromacra/Aesthetics), Kevin W. Boyack (SciTech Strategies, Inc.), Sara Irina Fabrikant (Geography Department, University of Zurich, Switzerland), Peter A. Hook (Law Libraries, Indiana University), Andri Skapin (Geography, San Diego State University), Bonnie DeGroot (Routledge), and Dave Whitt (Geography and Oceanography, Oregon State University). External reports that reviewed iterations 4 through 7 included: John R. Hiber (Chief of the Geography and Map Division, Library of Congress), Thomas B. Hickley (OLCL), Michael Kurtz (Harvard-Smithsonian Center for Astrophysics), Denise A. Bedford (World Bank), William Ying (CIO, ArtSTOR), Michael Korn (JSTOR), Carl Lagoze (Cornell University), Richard Furuta (Florida A&M University), Vincent Larivière (Université du Quebec à Montreal, Canada), Adam Bly (Idea: The New York Times), Alex Whitt (author of Guide: Mastering Information Through Online Agents), and Mills Norman (ProjectX.com).

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,” “Visualizing Science & Technology,” and “Fog and Fly Macrosopes” (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 write papers or otherwise cite these books in academic circles. And yet it seems quite natural to engage those who finance science via tax dollars to benefit from the research results—for bringing 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 researchers of the sea; David Attenborough, especially with his Life on Earth and Living Planet series; Paul Ott, with his Universal Atlas or Encyclopaedia Universalis Mammalorum; Stuart Brand, author of The Whole Earth Catalog; Richard Dawkins, famed for his Gombrich Up in the Umbrage; Al Gore for his environmental efforts, as featured in the Smithsonian’s First documentary and Hans Rosling, whose Gapminder effort gave rise to the metric, “Let my data change your mind.” It is my hope that this atlas serves as a springboard for both inspiration and encouragement to future science communicators.

Coping with the atlas was performed by Gordan Jelisavcevic, atlas design and layout by Tracy Thurtell, with many of the images specifically created for this book by Pratai Mavou-Lujan and Samuel T. Mills. Reference checks and formatting by Todd N. Thurtell, and copyright acquisition by Samantha Hale, Brunnah Marshall, Joseph Steinkopf, and Michael P. Gielo. Other valued contributions are acknowledged in the References & Credits (page 170).

This atlas was influenced by research and developments in many areas of science; it also benefited from countless discussions and brainstorming sessions with esteemed colleagues. And yet the brainstorming 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 narrating and thought provoking family dinner discussions and empowering girls’ night out. My gratitude also goes with our cat, Fiji, who kept me company through the many long periods of writing.

https://mitpress.mit.edu/books/atlas-forecasts
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>
Modeling Goals

Models aim to capture key phenomena at the levels that are most relevant for the understanding, communication, and management of systems. This spread describes and exemplifies key phenomena that are commonly studied when aiming to understand complex systems. Phenomena are roughly organised by question type (temporal, spatial, topical, and network) and complexity. Models that are static reference systems and no feedback cycles are introduced first, followed by phenomena that aim to capture evolving networks and activity patterns unfolding over them, including feedback or causal loops.

The great churning of the human race is our inability to understand the underlying forces.

—Albert H. Blakeslee

Phenomena of Interest

[i]ni the case of other phenomena that model and are characterised in a target system and are comprehensively defined above as a system of states to capture. For any modeling effort, it should start with identification of the phenomena to be included, together with information on system-specific simplifications that might or might not be acceptable. These simplifications can be readily achieved models that contain some values (see [Chatten et al. on page 24]). A model might have various states to observe particular properties of phenomena (e.g., temporal, spatial, or categorical), which are often presented as the system’s output (see [Chatten et al. on page 24]). The output is often plotted to represent the system’s output (see [Chatten et al. on page 24]).

Tipping Point

A tipping point is a critical threshold, a critical change point when small changes make a critical difference. Tipping points occur when a system changes from one state to another, and these changes can be hysteresis phenomena. Tipping points are also referred to as critical phenomena.

In 2007, for example, the tipping point of climate change was reached when the temperature of the earth exceeded the threshold of the ice age, resulting in the current global warming trend. Tipping points are also observed in ecology, where the loss of a key species can cause a cascade of changes, leading to the collapse of an ecosystem.

Reaction-Diffusion Dynamics

The reaction-diffusion model is a mathematical model used to describe the spread of phenomena in a system. It is based on the idea of a reaction-diffusion equation, which describes how the concentration of a substance changes over time and space. The reaction-diffusion model is used in a wide range of applications, from the spread of diseases to the patterns formed in chemical reactions.

The reaction-diffusion model is a mathematical model used to describe the spread of phenomena in a system. It is based on the idea of a reaction-diffusion equation, which describes how the concentration of a substance changes over time and space. The reaction-diffusion model is used in a wide range of applications, from the spread of diseases to the patterns formed in chemical reactions.

Phenomena

Oscillation

Phase Transition

Synchronization

Perspective

Phase Transition

The transformation of a thermodynamic system from one phase of state to another is called a phase transition. Examples include the melting of ice, the boiling of water, and the condensation of water vapor. Phase transitions are important in many fields, including physics, chemistry, and biology. They are also used in the development of new materials, such as superconductors and semiconductors.

Perspective

Phase Transition

The transformation of a thermodynamic system from one phase of state to another is called a phase transition. Examples include the melting of ice, the boiling of water, and the condensation of water vapor. Phase transitions are important in many fields, including physics, chemistry, and biology. They are also used in the development of new materials, such as superconductors and semiconductors.

Perspective

Phase Transition

The transformation of a thermodynamic system from one phase of state to another is called a phase transition. Examples include the melting of ice, the boiling of water, and the condensation of water vapor. Phase transitions are important in many fields, including physics, chemistry, and biology. They are also used in the development of new materials, such as superconductors and semiconductors.
Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding the model components, their behavior, the environment, and system dynamics evolving over time. Any model design 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 requirements must be detailed. Different approaches have been proposed to provide templates and standards for systematic model development and documentation—to ensure the replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data-driven framework as a supplement to the Atlas of Knowledge. Part 1, to cover the emerging phenomena discussed in the prior specialty, as well as the extended, descriptive, and predictive models discussed throughout the Atlas of Phenomena. We cannot stop the march of history, but we can influence its direction.

Visual Climate and adjacent components, both implicit and explicit, are the fundamental components of a metamodel. A metamodel is a starting point of a "modeling life's" to the various models.

Process

1. Analyze the existing information and data on the subject.
2. Define the scope and objectives of the modeling effort.
3. Choose the appropriate modeling method or technique.
4. Develop the model structure and framework.
5. Implement the model components and relationships.
6. Validate and verify the model.
7. Document the results and findings.

Graphic Symbol Types

- Geometric Symbol Types
- Line Symbol Types
- Linear Symbol Types
- Rectangular Symbol Types

Metamodel

A metamodel is a conceptual representation of a class of models and is used to define the structure of a modeling language. It specifies the elements that can be used in a model and the relationships between them. A metamodel is a blueprint for a modeling language that can be used to create models that conform to the language.

Modeling effort 2 is essential for the design and development of models. It is important to ensure that the final model, whether it is a simplified or a complex representation, is accurate and reliable. The process of modeling involves the following steps:

1. Define the problem: Identify the problem to be solved and the goals and objectives.
2. Model selection: Choose the appropriate modeling technique or method.
3. Model development: Build the model based on the chosen technique.
4. Model validation: Test the model to ensure its accuracy and reliability.
5. Model application: Use the model to solve the problem and make decisions.

16 Part 2: Methods

17 Part 2: Methods
Model Visualization

Model assumptions, designs, 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 clearly 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. Further VISUALIZATIONS may help stakeholders compare and interpret model results, and thus communicate their expertise to general audiances. VISUALIZATIONS can be static, dynamic, or interactive.


Visualization Types

The design of effective visualizations requires identifying sought needs and problems, selecting the appropriate data set, modeling, and visualizing types and performing an in-depth analysis of these variables in graphic symbols, as well as to visualize more complex, structural elements. This can involve the translation and encoding elements into visual codes. A number of existing models have been extended in to Modeling Framework (page 8).

Basic Model (Magnetic Model)

Model setup, when involving a large number of interacting information, the visualization and feature visualization of analysis, which are not easy to understand, can be broken down to tabular, time-series, or the like. The visualization can therefore be a graph, a number, a chart, or a table, which is easy to interpret.

The visualization can be a single or multiple-visualization, a graph or a visual object, a pattern, a structure, or a model. A single-visualization can be represented by a chart, a graph, a table, an image, a video, a sound, or a combination of these. A multiple-visualization can be represented by a set of charts, graphs, tables, images, videos, sounds, or a combination of these.

Model Conceptualization

The low-level pattern is that a model that conceptualizes all the details of a system, such as rates, states, variables, and so on. Different types of visualizations can be used to support this task.

The high-level pattern is that a model that conceptualizes all the details of a system, such as rates, states, variables, and so on. Different types of visualizations can be used to support this task.

10 Part 2 | Methods

28 Part 2 | Methods

Table and bar chart with a line chart and a pie chart. The line chart shows the trend over time, while the bar chart compares the values of different categories.

For example, the pie chart shows the proportion of the budget allocated to different departments in a company. The line chart can illustrate how the budget has changed over time. The bar chart can compare the budget allocations between different years.
Model Validation

Models should aim to capture the behavior of real-world systems in a simple yet useful 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 model (e.g., oscillations or avalanches) must reflect the phenomena observed in the real world. Models must be evaluated based on the accuracy and generality of their predictions. Evaluation results need to be used to increase the accuracy, specificity, or generalizability of the model, or to make model results easier to understand and to judge by decision-makers.

The more quantitative social indicator used for social decision-making, the more subject it will be to corruption, the more open it should be to adversarial and counterfactual analysis (it is intended to monitor).

Donald C. Campbell

Quality Assurance Framework

Models are subject to errors and imperfections that cannot exist in the real world. Models can mirror an observed event, but they cannot exactly match. Models can mirror an event and be useful, but they cannot be perfect. Models are used as an approximation, not an exact representation of the real world.

Consequences of approximating systems models
designed
to support target system simulation, decision-
making, and policy analysis. Here we describe how,
for a formal model—and that model should be
defined as a system model should be used to
design, simulation, and computer đẹp to identify
the stakeholder in the simulation, the model
representation of the system, and the computer
environment in which the model is used. This
means that the model is used to identify and
simulate the system, and the computer is used
to support simulation, and decision.

Consequence: the model is defined as a
design tool that is used to support
target system simulation, decision-
making, and policy analysis. Here we describe
how, for a formal model—and that model should be
defined as a system model should be used to
design, simulation, and computer đẹp to identify
the stakeholder in the simulation, the model
representation of the system, and the computer
evironment in which the model is used. This
means that the model is used to identify and
simulate the system, and the computer is used
to support simulation, and decision.

The major purpose of this classification is to
determine the potential consequences of using
systems models that are used to support
target system simulation, decision-making,
and policy analysis. Here we describe how,
for a formal model—and that model should be
defined as a system model should be used to
design, simulation, and computer đẹp to identify
the stakeholder in the simulation, the model
representation of the system, and the computer
evironment in which the model is used. This
means that the model is used to identify and
simulate the system, and the computer is used
to support simulation, and decision.

Model Verification

Model verification is the process of determining whether a computer model is a reasonable representation of its real-world counterpart. Model verification involves a systematic approach to testing the model and its assumptions. For example, a model of a chemical reaction might be tested by comparing its predictions to experimental data. If the model predictions are close to the experimental data, the model is considered to be verified. If the model predictions are not close to the experimental data, the model is considered to be invalid.

Model validation is the process of determining whether a computer model is a reasonable representation of its real-world counterpart. Model validation involves a systematic approach to testing the model and its assumptions. For example, a model of a chemical reaction might be tested by comparing its predictions to experimental data. If the model predictions are close to the experimental data, the model is considered to be validated. If the model predictions are not close to the experimental data, the model is considered to be invalid.

Model Replication

Model replication involves translating the model into a physical or computational system that can be used to test the model's predictions. For example, a model of a chemical reaction might be translated into a physical system that can be used to test the model's predictions. If the physical system produces results that are similar to the model's predictions, the model is considered to be replicated. If the physical system produces results that are different from the model's predictions, the model is considered to be invalid.

Model Reproducibility

Model reproducibility involves the ability to reproduce the model's predictions using the same or similar methods. For example, a model of a chemical reaction might be reproduced using the same or similar methods. If the reproduced predictions are similar to the original predictions, the model is considered to be reproducible. If the reproduced predictions are different from the original predictions, the model is considered to be un reproducible.

QA at Different Model Stages

QA is performed in different stages of a model's development. Here we describe how QA is performed in different stages of a model's development. For example, QA is performed in different stages of a model's development. QA is performed in different stages of a model's development.

QA at Different Model Stages

QA is performed in different stages of a model's development. Here we describe how QA is performed in different stages of a model's development. For example, QA is performed in different stages of a model's development. QA is performed in different stages of a model's development.

QA at Different Model Stages

QA is performed in different stages of a model's development. Here we describe how QA is performed in different stages of a model's development. For example, QA is performed in different stages of a model's development. QA is performed in different stages of a model's development.

QA at Different Model Stages

QA is performed in different stages of a model's development. Here we describe how QA is performed in different stages of a model's development. For example, QA is performed in different stages of a model's development. QA is performed in different stages of a model's development.

QA at Different Model Stages

QA is performed in different stages of a model's development. Here we describe how QA is performed in different stages of a model's development. For example, QA is performed in different stages of a model's development. QA is performed in different stages of a model's development.

Model Comparison

Model comparison involves comparing the results of different models to determine which model is more accurate or which model is more appropriate for a particular application. For example, a model of a chemical reaction might be compared to another model of a chemical reaction. If the two models produce similar results, the model is considered to be comparable. If the two models produce different results, the model is considered to be uncomparable.
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 phenomena such as artificial life, plant growth, or epidemiology. CAs consists of elements called cells. Each cell has a state, and cells are connected to form a one- or multi-dimensional lattice. Cell states change at discrete time steps 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 to implement self-replicating organisms. In 1968, Wolfram published a Basic Model Similar to Conway’s Game of Life (GOL) in the appendix of a paper. Later, cellular automata became a popular modeling framework for simulating complex systems and are widely used in fields such as physics, biology, and computer science.

In this discrete time step, cells are updated according to predefined rules based on their state and the states of their neighbors. The rules are applied to each cell independently, and the new state of the cell is determined by the current state of the cell and the states of its neighbors.

**Rule 30**

Rule 30, known as the majority rule, creates a chaotic pattern. In each step, each cell is updated based on the majority state of its neighbors. If the majority state is 1, the cell is updated to 1; otherwise, it is updated to 0. The pattern created by Rule 30 is known for its randomness and complexity.

**Terminology**

- **Tetrahedron**: The simplest cellular automaton, consisting of four cells arranged in a tetrahedral shape.
- **Hexahedron**: A cellular automaton consisting of six cells arranged in a hexagonal shape.
- **Systole**: A cellular automaton consisting of nine cells arranged in a systolic pattern.
- **Neuron**: A cellular automaton consisting of 27 cells arranged in a neuron-like pattern.

**Key Insights**

- CAs are used in various fields for modeling phenomena such as fluid dynamics, chemical reactions, and biological systems.
- CAs can be used to study complex systems and predict emergent behavior under certain conditions.
- CAs are useful for simulating natural and artificial processes, such as the growth of tissues, the spread of diseases, and the behavior of fluids.

**Conway’s Game of Life**

In the late 1960s, British mathematician John H. Conway invented the Game of Life, which was later popularized by mathematician John Horton Conway. The game is a 2D cellular automaton that exhibits a variety of complex behaviors, including self-replication, oscillation, and the formation of patterns that resemble gliders or spaceships.

**Schelling’s Segregation Model (1971)**

In 1971, the sociologist Thomas C. Schelling showed that individual bias can lead to collective behavior. His work was inspired by the Civil Rights Act of 1960, which required that “all persons—white, Negro, or Oriental—are alike, and they are not a random mixture.”

- **Phenomena**: People choose neighbors of the same or nearly the same race.
- **Mechanism**: Individuals choose to live with others of similar race, even if they are not a random mixture.
- **Conclusion**: Schelling’s model suggests that small biases can lead to large-scale segregation and social segregation.

**Schelling’s Model**

In his model, each agent is represented as a circle with a radius 0.5 units. The circle is placed on a grid, and each agent is assigned a race (black or white). The model shows how small biases can lead to large-scale segregation, even if the biases are not strong enough to cause segregation on their own.
Geospatial models—"Where"

Geospatial position and vectors are significant. Location data is often linked to vehicles, homes, or other locations, and analyzing this data can be useful for various applications. Geospatial models can help answer questions like: where is this vehicle located? or what is the closest location to a certain point?

Topical Models—"What"

Topical models are used to analyze information related to a specific topic. These models can help answer questions like: what are the key topics discussed in this document? or what are the most common words associated with a particular subject?

Network Models—"With What"

Network models are used to analyze relationships between entities. These models can help answer questions like: what entities are connected in this network? or what are the most important entities in this network?

Historical Progression

The evolution of network models has been driven by advancements in technology, computational power, and data availability. From simple models like graph theory to complex models like neural networks, the field of network models has grown significantly.

Network topology and node centrality are important factors in understanding the structure and function of networks. These concepts help answer questions like: what is the most central node in this network? or how is information spread through the network?
Reducing Human Bias

Humans tend to be subjective, often acting according to biased opinions rather than objective logic. Cognitive biases are systematic deviations from normative rationality in judgment, as evidenced in fields like psychology and behavioral economics. While many such biases have been confirmed independently reproducible research, many researchers are left asking how to make objective, well-informed decisions. This requires us to understand and proactively address our own biasing effects. This paper aims to explain human biases, their underlying mechanisms, and behaviors, with suggestions on how to counteract them. Ultimately, biases and beliefs have a major impact on life satisfaction. Understanding our own biases is an important step in constructing a fulfilling present and future.

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

To Erle Is Human

Though human beings are demonstrably more advanced and evolved, there are various instances at any level of resolution where human error in judgment and actions can be observed. The concept of human error in judgment is often used to reflect on the behavioral and decision-making processes of humans. This phenomenon is a well-documented aspect of human behavior, affecting personal, social, and economic outcomes. The study of human error in judgment is critical to understanding the limits and capabilities of human cognition.

Data Bias

Any system modeling or creating data, which is generated by sensing human interactions, can introduce errors or failures. Such errors or failures can affect the system’s output, leading to incorrect decisions or actions. These errors can be categorized based on various factors such as the reliability of the data, the complexity of the system, or the level of human intervention. Understanding and mitigating data bias is crucial to ensuring the accuracy and reliability of the system.

Gender Bias

Gender bias can be defined as the unfair treatment of individuals based on their gender. It is a form of discrimination that can manifest in various ways, including unequal pay, limited opportunities, and gender stereotypes. Understanding gender bias is essential to creating a more equitable society, where everyone has the opportunity to reach their full potential.

Self-Perpetuating Bias

Self-perpetuating bias refers to the tendency of a system to reinforce an existing belief or behavior, even when it is not beneficial. This can occur in various contexts, including scientific research, economic models, and social systems. Understanding self-perpetuating bias is crucial to developing more effective and equitable decision-making strategies.

Other Species

While humans have unique biases, other species also exhibit biases. Understanding animal biases can provide valuable insights into the evolution of human cognition and decision-making processes. This can help researchers better understand the underlying mechanisms of bias and develop more effective strategies to mitigate them.

Exposing Biases

People tend to resist changes to their own biases and often perceive other people as biased. However, exposing biases can help individuals recognize and address their own biases, leading to improved decision-making and increased productivity. This can be achieved through various methods, including self-reflection, feedback sessions, and exposure to diverse perspectives.

Part 5: Envisioning Desirable Futures

Modeling Opportunities

Reducing Human Bias

Managing Risks

Building Capacity

Actionable Forecasts
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 Börner, Filipi Nascimento Silva and Staša Milojević

Abstract | The numbers, size and complexity of ‘big science’ projects are growing— as are the time, space and data costs of big science projects. In this context, big data brings a new way to analyse, understand and communicate the inner workings of collaborations that often involve thousands of experts, hundreds of scholarly publications, thousands of new instruments and petabytes of data. The evolving topological and conceptual landscape of big science projects in physics, astronomy and biomedical sciences is a major contribution to the emerging science of knowledge production and use. In this context, big data can be a powerful tool for understanding the nature of big science, revealing patterns and trends that are not immediately apparent in the raw data. The use of big data analytics and visualization tools can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

Big science projects are often characterized by large, complex and dynamic networks of collaboration. These networks can be difficult to visualize and understand, especially when dealing with large datasets and complex interactions. However, by leveraging the power of big data analytics and visualization tools, we can gain valuable insights into the structure, dynamics and impact of these networks. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

The visualization of big science projects is a critical aspect of understanding the nature of these complex systems. By using visualization tools, we can gain a better understanding of the relationships between different data sets and the interactions between different actors. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

Big science projects are characterized by large, complex and dynamic networks of collaboration. These networks can be difficult to visualize and understand, especially when dealing with large datasets and complex interactions. However, by leveraging the power of big data analytics and visualization tools, we can gain valuable insights into the structure, dynamics and impact of these networks. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

The visualization of big science projects is a critical aspect of understanding the nature of these complex systems. By using visualization tools, we can gain a better understanding of the relationships between different data sets and the interactions between different actors. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

Big science projects are characterized by large, complex and dynamic networks of collaboration. These networks can be difficult to visualize and understand, especially when dealing with large datasets and complex interactions. However, by leveraging the power of big data analytics and visualization tools, we can gain valuable insights into the structure, dynamics and impact of these networks. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

The visualization of big science projects is a critical aspect of understanding the nature of these complex systems. By using visualization tools, we can gain a better understanding of the relationships between different data sets and the interactions between different actors. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

Big science projects are characterized by large, complex and dynamic networks of collaboration. These networks can be difficult to visualize and understand, especially when dealing with large datasets and complex interactions. However, by leveraging the power of big data analytics and visualization tools, we can gain valuable insights into the structure, dynamics and impact of these networks. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

The visualization of big science projects is a critical aspect of understanding the nature of these complex systems. By using visualization tools, we can gain a better understanding of the relationships between different data sets and the interactions between different actors. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

Big science projects are characterized by large, complex and dynamic networks of collaboration. These networks can be difficult to visualize and understand, especially when dealing with large datasets and complex interactions. However, by leveraging the power of big data analytics and visualization tools, we can gain valuable insights into the structure, dynamics and impact of these networks. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.

The visualization of big science projects is a critical aspect of understanding the nature of these complex systems. By using visualization tools, we can gain a better understanding of the relationships between different data sets and the interactions between different actors. This can help to identify key players, collaborations and trends in big science projects, providing insights into the underlying structures and dynamics of these complex systems.
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.

---

**Insight Needs**
- categorize/clustering
- order/hierarchies
- 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**
- point
- line area
- surface volume
- linguistic symbols text
- numerals
- punctuation marks
- pictorial symbols images
- icons
- statistical glyphs

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

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

---

1. **Interpret**
2. **Acquire**
3. **Analyze**
4. **Visualize**
5. **Stakeholders**
6. **Deploy**
Data Visualization Literacy Framework (DVL-FW)

Consists of two parts that are interlinked:

DVL Typology + DVL Workflow Process
Visual Analytics Certificate -

Interpret Deploy

Acquire Analyze Visualize

Stakeholders

Insight Need Types

Interaction Types

Data Scale Types

Analysis Types

Graphic Variable Types

Graphic Symbol Types

Visualization Types

1 2 3 4 5 6 7
Visual Analytics Certificate -

Interpret Deploy

Acquire Analyze Visualize

Stakeholders

Interaction Types

Data Scale Types

Insight Need Types

Operationalize

Translate

Design Data Overlay

Pick Reference System

Visualization Types

Graphic Symbol Types

Graphic Variable Types

Data Scale Types

Analysis Types

Pick Reference System
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

<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>• search and locate</td>
<td>• filter</td>
</tr>
<tr>
<td>• distributions (also outliers, gaps)</td>
<td>• interval</td>
<td>• geospatial</td>
<td>• graph</td>
<td>line</td>
<td>• details-on-demand</td>
<td>• history</td>
</tr>
<tr>
<td>• comparisons</td>
<td>• ratio</td>
<td>• topical</td>
<td>• map</td>
<td>area</td>
<td>• extract</td>
<td>• extract</td>
</tr>
<tr>
<td>• trends (process and time)</td>
<td></td>
<td>• relational</td>
<td>• tree</td>
<td>surface</td>
<td>• link and brush</td>
<td>• projection</td>
</tr>
<tr>
<td>• geospatial</td>
<td></td>
<td></td>
<td>• network</td>
<td>volume</td>
<td>• distortion</td>
<td></td>
</tr>
</tbody>
</table>
Typology of the Data Visualization Literacy Framework

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

- **Data Scales**
  - nominal
  - ordinal
  - interval
  - ratio

- **Analyses**
  - statistical
  - temporal
  - geospatial
  - topical
  - relational

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

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

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

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

<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></td>
<td></td>
<td>categorize/cluster</td>
</tr>
<tr>
<td>order</td>
<td>rank</td>
<td>ranking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>order/rank/sort</td>
</tr>
<tr>
<td>distribution</td>
<td>distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distributions (also outliers, gaps)</td>
</tr>
<tr>
<td>compare</td>
<td>nominal</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>time series</td>
<td>time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>track rises and falls over time</td>
</tr>
<tr>
<td>geospatial</td>
<td>geospatial</td>
<td>location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>geospatial</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>relations 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</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></td>
<td>motion</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

- Scatter Graph
- Geospatial Map
- UCSD Science Map
- Network

Graphic Symbols

Graphic Variables

Visualizations

CNS Center for Network Science
# 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 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>form</td>
<td>filter</td>
</tr>
<tr>
<td>comparisons</td>
<td>ratio</td>
<td>topical</td>
<td>map</td>
<td>area</td>
<td>color</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>optics</td>
<td>history</td>
</tr>
<tr>
<td>geospatial</td>
<td></td>
<td></td>
<td>network</td>
<td>volume</td>
<td>motion</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>

## 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>• search and locate</td>
<td>• filter</td>
</tr>
<tr>
<td>• distributions (also outliers, gaps)</td>
<td>• interval</td>
<td>• geospatial</td>
<td>• graph</td>
<td>line</td>
<td>• details-on-demand</td>
<td>• extract</td>
</tr>
<tr>
<td>• comparisons</td>
<td>• ratio</td>
<td>• topical</td>
<td>• map</td>
<td>area</td>
<td>• link and brush</td>
<td>• projection</td>
</tr>
<tr>
<td>• trends (process and time)</td>
<td></td>
<td>• relational</td>
<td>• tree</td>
<td>surface</td>
<td>• distortion</td>
<td>• correlation</td>
</tr>
<tr>
<td>• geospatial</td>
<td></td>
<td></td>
<td>• network</td>
<td>volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• compositions (also of text)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• correlations/relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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

#### Geometric Symbols

<table>
<thead>
<tr>
<th>Spatial Position</th>
<th>Geometric Symbols</th>
<th>Linguistic Symbols</th>
<th>Pictorial Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point</td>
<td><img src="image1.png" alt="Point Symbols" /></td>
<td><img src="image2.png" alt="Text Symbols" /></td>
<td><img src="image3.png" alt="Pictorial Symbols" /></td>
</tr>
<tr>
<td>Line</td>
<td><img src="image4.png" alt="Line Symbols" /></td>
<td><img src="image5.png" alt="Text Symbols" /></td>
<td><img src="image6.png" alt="Pictorial Symbols" /></td>
</tr>
</tbody>
</table>

#### Qualitative

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

#### Quantitative

- Also called: Ordered Attributes
- Also called: Magnitude Channels

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
Accelerating Behavioral Science Through Ontology Development and Use

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 entry type and their diverse interconnections.

This project is funded by NSF award 2033503.

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 interlinked 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+). 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 three-dimensional (3D) organization of whole organs and thousands of 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 funding from multiple sources. The 16 consortia include the Allen Brain Atlas1, the Brain Research through Advancing Innovative Neurotechnologies Initiative—Cell Census Network Initiative2, the Chan Zuckerberg Initiative Seed Networks for HCA3,4, HCA awards by the EU’s Horizon 2020 program, the Genotype-Tissue Expression project, the GenitoUrinary Developmental Molecular Anatomy Project5, Helmsley Charitable Trust: Gut Cell Atlas6,7,8, the Human Tumor Atlas Network9, the Human Biomedical Atlas Program (HuMap)10, the Kidney Precision Medicine Project (KPM)11,12,13, LungMAP14, HCA grants from the United Kingdom Research and Innovation Medical Research Council (https://mrc.ukri.org), (Re)building the Kidney15, Stimulating Peripheral Activity to Relieve Conditions16, The Cancer Genome Atlas17,18 and Wellcome funding for HCA pilot projects19,20. 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 coherent tomography (OCT)21.

https://www.nature.com/articles/s41556-021-00788-6

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; H2CO, Horizon 2020; GTRx, Genotype-Tissue Expression project; SUDMAP, GenitoUrinary Developmental Molecular Anatomy Project; HTGA, Human Tumor Atlas Network; MRC, Medical Research Council; RBK, (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 11 organs. c, An exemplary ASCt+ table showing anatomical structures (AS) and cell types (CT) and some biomarkers (B) for each cell type. The relationships are (1) interlink major anatomical structures, cell types and biomarkers and (4) four relationship types (part_of, location, and characterizes). Note that the terminology exists for cell types and biomarkers.
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/