Mapping 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

2022 Chengdu Conference on Scientometrics & Evaluation Chengdu, China

8:30pm ET on Nov. 30, 2022 (Washington DC) | 9:30am on Dec. 1, 2022 (Beijing)
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)
The Power of Maps

Iteration II (2006)
The Power of Reference Systems

Iteration III (2007)
The Power of Forecasts

Iteration IV (2008)
Science Maps for Economic Decision-Makers

Iteration V (2009)
Science Maps for Science-Policy Makers

Iteration VI (2010)
Science Maps for Scholars

Iteration VII (2011)
Science Maps as Visual Interfaces to Digital Libraries

Iteration VIII (2012)
Science Maps for IGS

Iteration IX (2013)
Science Maps Showing Trends and Dynamics

Iteration X (2014)
The Future of Science Mapping

Macrosopes

Iteration XI (2015)
Macroscopes for Interacting with Science

Iteration XII (2016)
Macroscopes for Playing with Scale

Iteration XIII (2017)
Macroscopes for Making Sense of Science

Iteration XIV (2018)
Macroscopes for Measuring our Well-being

http://scimaps.org

100
MAPS in large format, full color, and high resolution.

248
MAPMAKERS from fields as disparate as art, urban planning, engineering, and the history of science.

43
MACROSCOPE MAKERS including one whose job title is “Truth and Beauty Operator.”

20
MACROSCOPEs for touching all kinds of data.

382
DISPLAY VENUES from the Cannes Film Festival to the World Economic Forum.

354
PRESS ITEMS including articles in Nature, Science, USA Today, and Wired.
Map of Scientific Collaborations from 2005-2009

VII.6 Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012
A Topic Map of NIH Grants 2007

Bruce W. Herr II, Gully A.P.C. Burns, David Newman, and Edmund Talley (NIH)

The National Institutes of Health (NIH) is organized as a multitude of Institutes and Centers whose missions are primarily focused on distinct diseases. However, disease etiologies and therapies 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 devised an interactive topic map for navigating this landscape, online at www.nihtopicmap.org. Institute abbreviations can be found at www.nih.gov/idc.

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 map focused on cardiac vascular functions and dysfunctions. Cardiac failure (generally funded by NHLBI) is topically clustered next to stroke (NIHRE). These are the two major medical emergencies associated with ischemia, which results from restricted blood supply. Also located in this area are grants focused on Heart Failure (NHLBI) signaling, a major biochemical pathway for vasodilation, and grants on hemodynamics, Stagger Cell Disease, and myocarditis.

Neural Circuits Research

A map focused on neuronal circuits, which shows the identity of topics and NIH Institutes that fund research in this area, such as: Cold Temporal Regulations, primarily funded by NHLBI; Visual Preprocessing, primarily handled by NIDCD; and Glymphatics, primarily funded by NINDS. For site loading, see legend in the upper left tool.

National Cancer Institute (NCI)

TOP 10 TOPICS
1. Oncology Clinical Trials
2. Cancer Treatment
3. Cancer Therapy
4. Carcinogenesis
5. Risk Factor Analysis
6. Cancer Chemotherapy
7. Genetics
8. Epidemiology
9. Prognosis
10. Cancer Chemoprevention

National Institute of General Medical Sciences (NIGMS)

TOP 10 TOPICS
1. Receptor Organic Synthesis
2. X-ray Crystallography
3. Protein Analysis
4. Computational Models
5. Yeast Biology
6. Metabolism
7. Enzyme Mechanisms
8. Protein Complexes
9. Translational/Genetic Genetics
10. Cell Divisions

National Heart, Lung, and Blood Institute (NHLBI)

TOP 10 TOPICS
1. Heart Failure
2. Pulmonary Injury
3. Genetic Linkage Analysis
4. Cardiovascular Disease
5. Atherosclerosis
6. Hemostasis
7. Blood Pressure
8. Asthma Allergy Airways Disease
9. Gene Association
10. Lipoproteins

National Institute of Mental Health (NIMH)

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

The Structure of Science

The Social Sciences are the smallest and least well-defined of all the sciences. Psychology serves as the bridge between Medical Science, Social Sciences, and Natural Sciences as it is concerned with the human psyche.

Mathematics is our starting point, the pivot of all sciences. It lies at the outer edge of the map. It is the foundation upon which all other scientific disciplines are constructed. Mathematics provides the tools for gaining new knowledge in Mathematics and Physics. These three disciplines provide a great example of a three-dimensional structure of science interrelations to another. Physics through inductive disciplines. Although applied, these disciplines are highly concentrated with distinct bands of research that can be identified. Band-mates interdisciplinary research.

Research is highly concentrated in Physics and Chemistry. These disciplines have their own distinct bands of research that can be identified. The broadness of these bands indicates an extensive amount of interdisciplinary research. Physics and Chemistry are not as distinct as one might assume.

The Life Sciences, including Biology and Bioinformatics, are less concentrated than Chemistry or Physics. Branches of life sciences can be unevenly distributed in the larger areas of the Life Sciences, for instance between Biology and Bioinformatics, and between Biology and Environmental Science. Bioinformatics is a very new discipline in the field. It is a large discipline that has wide links to disciplinary areas of the map. The Life Sciences, including Biology, 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 created at the Institute for Scientific Information in 1983. It consisted of 41 areas of science from the citation patterns of 1,080 scientific papers. That map was large and confusing, but it didn’t convey enough of science to accurately define its structure.

Things are different today. We have enormous computing power and advanced visualization software that make sense of the structure of science possible. We have the map of science (1990) generated at Indiana University using an advanced graphical browser called VOSviewer from the citation patterns of 58,284 scientific papers published in 2003. This map is the result of analysis of the co-citation graphs of the 58,284 papers and the 96,000 research communities active in science in 2012. A research community is a group of papers that are written on the same research topic in a given year. Over time, communities can be born, continue, split, merge, or die. The map of science can be used as a tool for many different fields. It has the ability to inform the scientist and the researcher about the scientific and economic impact of their research community allows policy makers to decide which areas to expand, maintain, abandon, or ignore. We also calculate the map as an educational tool. For children, the hierarchical relationship between areas of science can be represented in a concrete map showing how math, physics, chemistry, biology and social sciences intermingle. For advanced students, areas of interest can be located and neighboring areas can be explored.

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

Proteomics
- Research communities in proteomics are concentrated proteomics. In addition, there is a heavy focus on the fields within of chemistry, such as Mass Spectrometry. The balance of the proteomics communities are typically centered among the Life and Medical Sciences.

Pharmacogenomics
- Pharmacogenomics is a relatively new field with most of its activity in Medicine. It also has many connections in both the natural sciences, except for the Social Sciences.

The Medical Sciences include broad therapeutic studies and targeted areas of treatment (e.g., centre nervous system, cancer, pharmacogenetics, etc.) Unlike Physics and Chemistry, the medical disciplines are more spread out, suggesting a more multidisciplinary approach to research. The transition into Life Sciences and related Science and Biochemistry is gradual.

Impact

The United States Patent and Trademark Office classifies and indexes patents using a hierarchical system of classifications. Patents are organized in a taxonomy that groups patents together into categories, subcategories, and more specific classifications. The hierarchical structure allows for the identification and retrieval of patents based on their content and similarity to other patents.

Taxonomy Visualization of Patent Data

- Katy Börner, Elisha F. Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley - 2006

The US Patent Hierarchy

Prior Art

New patents often build on older ideas from more different categories. How ideas flow through the system categories is the current patent visualizations. Patents are visualized in a way that shows how ideas have evolved over time. The visualizations highlight the relationships between patents, showing how ideas have been refined and expanded over the years.

Keeping categories understandable is an important part of maintaining any patent visualizations. Categories are crucial for categorizing and retrieving patents. They help users find relevant patents and understand the context in which they were developed. Interactive visualizations allow users to explore the relationships between categories and patents, providing a deeper understanding of the innovation landscape.
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/macroscope-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

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
MediX 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
- Macroscopes 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
Acknowledgments

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-1223898, OCE-0949245, SBE-0718112, and CBET-0831368; the National Institutes of Health under Grants No. U54 GM085599, R21-DH043259, and U34-RR029822; the Jones S. McDevitt 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 Scholar of Library and Information Science—all three 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 Wouters of CWTS, Andrea Scharnhorst and Jeannette Haagenaars of Mertens, and Peter Dresser, Linda Reijndorp, and Laura Patouميز 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 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 include: Deborah MacPherson (Accuracy/Aesthetics), Kevin W. Boyack (SciTech Strategies, Inc.), Sara Irina Fabriant (Geography Department, University of Zurich, Switzerland), Peter A. Hook (Law Libraries, Indiana University), Andri Stavrin (Geography, San Diego State University), Bonnie DiGrazi (Routledge, Ltd.), and Dana Wright (Geography and Oceanography, Oregon State University). External experts that reviewed iterations 4 through 7 included: John R. Hibert (Chief of the Geography and Map Division, Library of Congress), Thomas B. Hickey (OCLC), Michael Kurtz (Harvard-Smithsonian Center for Astrophysics), Denise A. Belford (World Bank), William Ying (CIO ArtSTOR), Michael King (JSTOR), Carl Logue (Cornell University), Richard Fortuna (Cies-McM University), Vincent Larivière (Université du Québec à Montréal, Canada), Adam Bly (CEO of WEED), Alice Wright (author of How Mastering Information Through Big Data, and Mills Chris (ProjectSylo.com)).

Focused brainstormsing workshops, organized with collaborators 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 Showcase Science,” “Mapping the History and Philosophy of Science,” “Modeling Knowledge Dynamics,” “Nurturing Evasion Science & Technology,” and “Flag-and-Flap 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 a target to enable those who finance science via tax dollars to benefit from the research results—forbitting the maximization of citation counts via the production of research papers. Many others have taken this route, including the following honorary mixins 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 Oliver, with his Universal Atlas on Zoology, Universal Meteorology; Stewart Brand, author of The Whole Earth Catalogue; Richard Dawkins, famed for his “Growing Up in the Universe” lecture; Al Gore for his environmental efforts, as featured in the acclaimed Plevi documentary; and Hans Rosling, whose Gapminder effort gave rise to the motto, “Let your data change your mind.” It is my hope that this atlas seeks roles in giving both inspiration and encouragement to future science communicators.

Copying of this atlas was performed by Gomina Jellinevic, atlas layout and design by Tracy Tebars, with many of the images specifically curated for this book by Perla Mano-Leijn and Samuel T. Mills. Reference checks and formatting by Todd N. Therriault, and copyright acquisition by Samantha Hale, Brannon Marshall, Joseph Stankiewicz, and Michael P. Ceja. 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 esteemed colleagues. And yet the hierarchy 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 sharing providing family dinner discussions and encouraging girl’s nights out. My gratitude also goes with our cat, Jfi, 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 organized by question type (temporal, spatial, topological, 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 them, including feedback or causal loops.

The greatest dancing of the human race is our inability to understand the opposite.

— Albert E. Barlow

Phenomena of Interest

Identifying what phenomena models need to characterize is a major step in formulating a model. It is often necessary to first understand the phenomena that are relevant to understanding a system. Phenomena can be divided into four major categories: Oscillation, Tipping Point, Synchronization, and Phase Transition. Each of these categories is described in detail below.

Oscillation

Any system that responds to a stimulus as an oscillation. Examples include a swinging pendulum, a biological oscillation, or a network oscillation.

Tipping Point

A point at which a system becomes unstable and changes from one state to another. Examples include a financial crisis, a phase transition, or a system collapse.

Synchronization

The phenomenon of two or more systems becoming synchronized in their behavior. Examples include the synchronization of heartbeats, the synchronization of fireflies, or the synchronization of biological rhythms.

Phase Transition

The transformation of a system from one state to another as a function of a parameter. Examples include the transformation of water from a liquid to a solid, the transformation of a chemical reaction from one state to another, or the transformation of a neural network from one state to another.

Adaptation & Learning

The process by which systems adapt and learn from their environment. Examples include the adaptation of a neuron to a new stimulus, the adaptation of a computer network to changes in traffic, or the adaptation of a chemical reaction to changes in temperature.

Fractal Dynamics

The study of systems that exhibit self-similarity across different scales. Examples include the fractal growth of a tree, the fractal distribution of stars in a galaxy, or the fractal patterns of a snowflake.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.

Target System Models

Pendulum

The classic example of a simple, single-degree-of-freedom system. Examples include a simple pendulum or a simple harmonic oscillator.

Coupled Oscillators

A system of two or more oscillators that interact with each other. Examples include a system of coupled pendulums or a system of coupled harmonic oscillators.

Networks

A system of interconnected nodes that interact with each other. Examples include a social network, a biological network, or a computer network.

Random Networks

A system of nodes that are connected randomly. Examples include a random network or a random graph.

Barabási Network

A network that exhibits a power-law distribution of node degrees. Examples include a social network or a biological network.

Scale-Free Network

A network that exhibits a power-law distribution of node degrees. Examples include a social network or a biological network.

Small-World Network

A network that exhibits a short average path length and a high clustering coefficient. Examples include a social network or a computer network.

Network Dynamics

The study of how networks evolve over time. Examples include the evolution of a social network, the evolution of a biological network, or the evolution of a computer network.

Evolutionary Dynamics

The study of how systems evolve over time. Examples include the evolution of a social network, the evolution of a biological network, or the evolution of a computer network.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.

Phase Transition

The transformation of a system from one state to another as a function of a parameter. Examples include the transformation of water from a liquid to a solid, the transformation of a chemical reaction from one state to another, or the transformation of a neural network from one state to another.

Adaptation & Learning

The process by which systems adapt and learn from their environment. Examples include the adaptation of a neuron to a new stimulus, the adaptation of a computer network to changes in traffic, or the adaptation of a chemical reaction to changes in temperature.

Fractal Dynamics

The study of systems that exhibit self-similarity across different scales. Examples include the fractal growth of a tree, the fractal distribution of stars in a galaxy, or the fractal patterns of a snowflake.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.

Target System Models

Pendulum

The classic example of a simple, single-degree-of-freedom system. Examples include a simple pendulum or a simple harmonic oscillator.

Coupled Oscillators

A system of two or more oscillators that interact with each other. Examples include a system of coupled pendulums or a system of coupled harmonic oscillators.

Networks

A system of interconnected nodes that interact with each other. Examples include a social network, a biological network, or a computer network.

Random Networks

A system of nodes that are connected randomly. Examples include a random network or a random graph.

Barabási Network

A network that exhibits a power-law distribution of node degrees. Examples include a social network or a biological network.

Scale-Free Network

A network that exhibits a power-law distribution of node degrees. Examples include a social network or a biological network.

Small-World Network

A network that exhibits a short average path length and a high clustering coefficient. Examples include a social network or a computer network.

Network Dynamics

The study of how networks evolve over time. Examples include the evolution of a social network, the evolution of a biological network, or the evolution of a computer network.

Evolutionary Dynamics

The study of how systems evolve over time. Examples include the evolution of a social network, the evolution of a biological network, or the evolution of a computer network.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.

Reaction-Diffusion Dynamics

The study of systems that exhibit wave-like patterns or oscillations. Examples include the Belousov-Zhabotinsky reaction, the Turing reaction, or the Belousov-Zhabotinsky reaction.
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 design should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define what a model is fit for purpose. Model validation and communication must be integrated. Different approaches have been proposed to provide guidelines and standards for systematic model development and documentation, which can improve the replicability of results. This special issue on modeling frameworks and other guidelines is an opportunity to draw on best practices, experiences, and lessons from previous modeling efforts.

Modeling Frameworks: A Call to Action

A modeling framework is a set of guidelines and principles that can help improve the quality and replicability of models. Frameworks provide a structure for organizing and communicating the components and processes of a model. They can help ensure that models are consistent, well-documented, and easier to understand and maintain.

Modeling Frameworks: A Taxonomy

Modeling frameworks can be classified into two main types: language frameworks and domain frameworks. Language frameworks focus on the syntax and semantics of the modeling language itself, providing a consistent way to represent models. Domain frameworks, on the other hand, focus on the representation of specific domains or systems, providing a way to model the specific phenomena and processes of interest.

Modeling Frameworks: A Call to Action

The special issue on modeling frameworks provides an opportunity to draw on best practices and lessons from previous modeling efforts. It is an opportunity to communicate model structure, dynamics, and results effectively across disciplines as well as to non-specialists—within academia, industry, and government policymaking. This special issue provides a platform for discussion and exchange of ideas in the modeling community and the modeling community itself.

Modeling Frameworks: A Call to Action

The special issue on modeling frameworks provides an opportunity to draw on best practices and lessons from previous modeling efforts. It is an opportunity to communicate model structure, dynamics, and results effectively across disciplines as well as to non-specialists—within academia, industry, and government policymaking. This special issue provides a platform for discussion and exchange of ideas in the modeling community and the modeling community itself.

Modeling Frameworks: A Call to Action

The special issue on modeling frameworks provides an opportunity to draw on best practices and lessons from previous modeling efforts. It is an opportunity to communicate model structure, dynamics, and results effectively across disciplines as well as to non-specialists—within academia, industry, and government policymaking. This special issue provides a platform for discussion and exchange of ideas in the modeling community and the modeling community itself.
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 a shared understanding 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 build communications to experts in general audiences. Visualizations can be static, dynamic, or interactive.

The beauty of visualization is simplicity. Chen, B., Walter, L., and Eisert, H.

Visualizations Types

The design of effective visualizations requires identifying image needs and goals; selecting the appropriate image type, modeling, visual design, and visualizations types; and performing an on-the-fly analysis of the types to validate and, as necessary to support visualizations 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 build communications to experts in general audiences. Visualizations can be static, dynamic, or interactive.

Simple model

The simple model provides a simple view of the 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 build communications to experts in general audiences. Visualizations can be static, dynamic, or interactive.

Model Conceptualization

Model Conceptualization

The simple model provides a simple view of the 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 build communications to experts in general audiences. Visualizations can be static, dynamic, or interactive.
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 model, oscillations or alternations, must reflect the phenomena observed in the real world. Models must be evaluated based on the accuracy and generality of their predictions. Evaluation tools need to be used to access the uncertainty, specificity, or generality of the model, or to make model results easier to understand and see by decision-makers.

More or any quantitative social indicator is used for social decision-making, the more subject is to an ever-growing proportion of the more is to be evaluated and-correct the social process it is intended to monitor.

Donald E. Campbell

Quality Assurance Framework

A quality assurance system (QAS) is one that helps enterprises understand any errors and classify, analyze, and monitor, and that is the subject of your project. The Quality Assurance Department of the organization (QA&D) evaluates the impact of changes, as well as updates its quality assurance systems by reviewing the framework for iterative changes. The framework for iterative changes begins with identifying the process of identifying and prioritizing feasible options and identifying the various components and elements of the model. A key component is the design of the model, which determines its function. A quality assurance framework is essential for understanding the components of a model, the processes it involves, and the elements it represents.

Model Precision and Accuracy

Accuracy refers to the precision of decision-makers to a standard or known real-time. Precise decisions are typically made on models that are as accurate as possible; in terms of an actual model, the accuracy is best when the model provides the actual model or the model that is closest to the actual model.

Model Robustness

The accuracy of the model is determined by measuring the changes in predictions and generates more varied input to the input data and process settings. Ideally, flexibility and uncertainty are added to the data model, and their impact on the model's results is considered. Model validation is used to determine the accuracy and robustness of the model, as well as to evaluate its validity.

For simple models with low data volume pre-approach

For complex models with high data volume pre-approach

Model Validation

Model validation is the process of determining whether an implemented method is a representative way to represent some phenomena in the real world. Model validation involves evaluating the reliability and validity of the model, its structure, and its results. A well-designed model should have been successfully implemented and is subject to scrutiny and criticism. Model validation is often performed after the model has been used for a sufficient number of iterations, and it is analyzed and discussed in detail.

Model Verification

Verification is the process of confirming that a model has produced the desired results. A model that has been validated is expected to have produced the desired results, and that model results are expected to be accurate. It is not always easy to verify or determine the accuracy of model results or observations at the cost of practicality.

Model Replication

Replication refers to the use of a model that is identical to the model used for the initial experiment. Model replication is used to determine the validity of the results obtained from the model, and to ensure that the results are not due to chance. Model replication is performed by using the same model parameters and initial conditions as the original model.

Model Comparison

Model comparison is the process of comparing the results of different models, which may be used to determine which model is the most accurate or which model is the best for a particular application. Model comparison is used to determine the relative performance of different models, and to determine which model is the most appropriate for a particular application.

Model Building QA

Developing models requires an understanding of the development of the model and how to extend any through the model.

Model Test and Deliver QA

Developing models requires an understanding of the development of the model and how to extend any through the model.

Review assumptions—Checking that assumptions are modeled using only the data available for data analysis.

Quality Assurance Framework (QAF) includes not only the characteristics that help define a particular model, but also the components that can be used to validate a model. The QAF system includes the following components:

Modeling and simulation are used to determine the accuracy and validity of the model. The accuracy of the model is determined by measuring the changes in predictions and generates more varied input to the input data and process settings. Ideally, flexibility and uncertainty are added to the data model, and their impact on the model's results is considered. Model validation is used to determine the accuracy and robustness of the model, as well as to evaluate its validity.
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 social networks—by statistical mechanisms. They include, for example, connected CA that describe the behavior of cells, and many other fields. CAs are used to model various phenomena, including the spread of diseases, the behavior of financial markets, and the evolution of ecosystems. In this context, the term “cellular automaton” refers to a mathematical model of a system that is composed of identical discrete elements, such as cells or particles, which interact with each other according to specific rules. The behavior of the system is then determined by the interactions between these elements. CAs can be used to simulate a wide range of phenomena, from simple patterns to complex systems. They are also used in computer science, particularly in the field of artificial intelligence, to study the behavior of complex systems and to design algorithms for machine learning. CAs are a powerful tool for understanding complex systems, and they are widely used in many fields of research.
The model questions overview.

Geospatial position and space are significant. Certain elements are bijective, with others either occurring in the overall system or interacting in the overall system. Therefore, the overall system is bijective, with others either occurring in the overall system or interacting in the overall system. Therefore, the overall system is bijective, with others either occurring in the overall system or interacting in the overall system. Therefore, the overall system is bijective, with others either occurring in the overall system or interacting in the overall system. Therefore, the overall system is bijective, with others either occurring in the overall system or interacting in the overall system.
Reducing Human Bias

Humans tend to be subjective, often acting according to biased opinions rather than objective facts. Cognitive biases are systematic deviations from normative rationality in judgment, as used in fields like psychology and behavioral economics. While many such biases have been confirmed in independently reproducible research, consensus about their cognitive basis is lacking. In order to make objective, well-informed decisions, we need to understand and proactively mitigate these biases. This includes understanding social biases, biases in decision-making, and behavioral, and cognitive biases, with suggestions on how to correct them. Ultimately, biases and beliefs have a major impact on life satisfaction. Understanding our biases and understanding how to correct them is an important first step in building a fulfilling present and future.

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

Reducing Human Bias

Humans tend to be subjective, often acting according to biased opinions rather than objective facts. Cognitive biases are systematic deviations from normative rationality in judgment, as used in fields like psychology and behavioral economics. While many such biases have been confirmed in independently reproducible research, consensus about their cognitive basis is lacking. In order to make objective, well-informed decisions, we need to understand and proactively mitigate these biases. This includes understanding social biases, biases in decision-making, and behavioral, and cognitive biases, with suggestions on how to correct them. Ultimately, biases and beliefs have a major impact on life satisfaction. Understanding our biases and understanding how to correct them is an important first step in building a fulfilling present and future.

All models are wrong, but some are useful.
George Box
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, with Filipi N. Silva and Staša Milojević, is out in @NatRevPhys, see rdcu.be /cyEG5. Explore interactive vis at bigscience.github.io then use code to map your very own projects.

@IUNetSci @IULuddy @cnscenter @ieeviis @issi_pres

PERSPECTIVES

Visualizing big science projects

Katy Börner, Filip Nascimento Silva and Staša Milojević

Abstract | The number, size and complexity of ‘big science’ projects are growing — as are the size, complexity and value of the data sets and software services they produce. In this context, big data gives a new way to analyse, understand, and manage and communicate the interests of large collaborations that often involve thousands of experts, thousands of scholarly publications, hundreds of new instruments and petabytes of data. We compare the evolving geographical and temporal impact of big science projects in physics, astronomy and biomedical sciences. A total of 13,893 publications and 1,139 grants by 21,945 authors cited more than 333,722 times are analysed and visualized to help characterize the distinct phases of big science projects, document increasing internationalization and densification of collaboration networks, and reveal the increase in interdisciplinary impact over time. All data sets and visual analytics workflows are freely available on GitHub in support of future big science studies.

‘Big science’ today is international, interdisciplinary and inter-institutional. Big science projects are anchored around expensive, large and complex instruments, they can run for several decades and they involve thousands of experts. Big science projects make breakthroughs not only in basic research but also in innovation that impacts economy and solves challenging societal needs. As more science fields move towards the big science model of knowledge creation, the lessons learned from previous successful endeavours become essential. This is described as ‘unleashing’ by classical philologist and Princeton Academy of Science member Theodor Mommsen: ‘The better known and more immediate precursors of what became known as big science are the establishment of the University of California cyclotron by Ernest Lawrence in the 1930s for energy research’ and the ‘World War II Manhattan Project’. The term ‘big science’, however, was introduced in the 1960s by Alan M. Weisberg5 and Derek J. De Solla Price to describe post-World War II developments in physics that built large and very expensive instruments (reactors and accelerators), accompanied by the growth in scientific teams working on nuclear-related research. Making advances in nuclear and, later, particle physics became part of the competition among superpowers, with the expectation that breakthroughs would lead to both scientific and technological superiority6,7. In addition, big science has been propelled into the general public awareness by the founding of the National Aeronautics and Space Administration (NASA) and its active and publicly visible space programmes8. Although most of the early focus regarding big science was on physics, as early as 1960, Weisberg9 proposed that biomedical science and biomedical technology were ready to enter the ‘big biology’ era. This entry was made only in the 1990s with the Human Genome Project (HGP), the first big science project in biology10. The expansion of the big science mode of knowledge production to other areas of science, such as big biology, brought with it new organizational and collaborative forms, such as ‘networked’ science enabled by information and communication technologies11 and some debates as to whether such coordinated efforts can be called big science12.

Big science accentuated the role of instruments in the development of science as engine of discovery13. Historically, instruments such as the telescope, the microscope and the air pump opened new vistas and led to scientific evolution, fundamentally changing the nature of scholarship14. The quest for increased sensitivity and accuracy of instruments led to their constant evolution, making these ever more expensive tools15,16 obsolete fairly quickly17. This process has been described as ‘unleashing’, in which ‘lineages of technology’ are adapted and combined, leading to networks, or ‘epistemic networks’, of technologies. However, the power of instruments, such as a scanning transmission electron microscope, can only be realized when they engage a community of researchers in what has been called ‘in situ’ experiments. ‘In situ’ is eventually leading to the formation of new scientific fields, such as nanotechnology. Furthermore, the relationship between science and technology is complex and interdependent, with science also contributing to technology development18. Early scientists, such as Galileo Galilei and Isaac Newton, engage in instrument building as well as theoretical and experimental work19. While not without precedent, instrument building
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://24hourssciencemap.info

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 (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.
Data Visualization Literacy Framework (DVL-FW)

Consists of two parts that are interlinked:

DVL Typology + DVL Workflow Process

1. Stakeholders
2. Data Scale Types
3. Analysis Types
4. Visualization Types
5. Graphic Symbol Types
6. Graphic Variable Types
7. Interaction Types
Visual Analytics Certificate -

Interpret Deploy
Acquire Analyze Visualize

Stakeholders
Insight Need Types

1 Data Scale Types
2 Analysis Types

Interaction Types

3

4 Visualization Types

Graphic Symbol Types

5

6 Graphic Variable Types
Visual Analytics Certificate -

Interpret

Deploy

Interpret

Stakeholders

Translate

Insight

Need

Types

Operationalize

Data Scale Types

Acquire

Analyze

Analysis Types

Visualize

Design

Data

Overlay

Pick

Reference

System

Graphic Variable Types

Graphic Symbol Types

Visualization Types
Data Visualization Literacy Framework (DVL-FW)

Implemented in Make-A-Vis (MAV) to support learning via horizontal transfer, scaffolding, hands-on learning, etc.
Typology of the Data Visualization Literacy Framework

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

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

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

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

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

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

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

## 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 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>• linguistic symbols</td>
<td></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>

<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></td>
<td>category</td>
<td></td>
<td></td>
<td>categorize/cluster</td>
<td>order/rank/sort</td>
<td>distributions (also outliers, gaps)</td>
</tr>
<tr>
<td>order</td>
<td>rank</td>
<td>ranking</td>
<td></td>
<td></td>
<td></td>
<td>table</td>
<td></td>
<td></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</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></td>
<td>&amp; deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<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>
</tr>
<tr>
<td>geospatial</td>
<td>spatial relations</td>
<td>location</td>
<td></td>
<td></td>
<td>generate maps</td>
<td>geospatial</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</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</td>
<td></td>
<td>• projection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• numerals</td>
<td></td>
<td>• distortion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• punctuation marks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• pictorial symbols</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• images</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• icons</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• statistical glyphs</td>
<td></td>
<td></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

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

**Table**
columns by rows

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>row</td>
<td>column</td>
</tr>
<tr>
<td></td>
<td>cell</td>
</tr>
</tbody>
</table>

**Graph**
x-y coordinates

Y

\[ X \]

**Map**
latitude/longitude

Y

\[ longitude \]

**Network**
local similarity

Y

\[ X \]

\[ node \]

\[ edge \]
Visualize: Reference Systems, Graphic Symbols and Variables

Reference System

Scatter Graph  Geospatial Map  UCSD Science Map  Network

Graphic Symbols

Graphic Variables

Visualizations
## 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

---

Typology of the Data Visualization Literacy Framework

**Insight Needs**
- categorize/clustering
- 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

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>Graphic Variable Types</th>
<th>Geometric Symbols</th>
<th>Linguistic Symbols</th>
<th>Pictorial Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial Position</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>![X Symbol]</td>
<td>![X Symbol]</td>
<td>![X Symbol]</td>
</tr>
<tr>
<td>Y</td>
<td>![Y Symbol]</td>
<td>![Y Symbol]</td>
<td>![Y Symbol]</td>
</tr>
<tr>
<td><strong>Form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>![Size Symbol]</td>
<td>![Size Symbol]</td>
<td>![Size Symbol]</td>
</tr>
<tr>
<td>Shape</td>
<td>![Shape Symbol]</td>
<td>![Shape Symbol]</td>
<td>![Shape Symbol]</td>
</tr>
<tr>
<td><strong>Value</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Color</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hue</td>
<td>![Hue Symbol]</td>
<td>![Hue Symbol]</td>
<td>![Hue Symbol]</td>
</tr>
<tr>
<td>Saturation</td>
<td>![Saturation Symbol]</td>
<td>![Saturation Symbol]</td>
<td>![Saturation Symbol]</td>
</tr>
<tr>
<td>Granularity</td>
<td>![Granularity Symbol]</td>
<td>![Granularity Symbol]</td>
<td>![Granularity Symbol]</td>
</tr>
<tr>
<td>Texture</td>
<td>![Texture Symbol]</td>
<td>![Texture Symbol]</td>
<td>![Texture Symbol]</td>
</tr>
<tr>
<td>Pattern</td>
<td>![Pattern Symbol]</td>
<td>![Pattern Symbol]</td>
<td>![Pattern Symbol]</td>
</tr>
<tr>
<td><strong>Blur</strong></td>
<td>![Blur Symbol]</td>
<td>![Blur Symbol]</td>
<td>![Blur Symbol]</td>
</tr>
<tr>
<td><strong>Speed</strong></td>
<td>![Speed Symbol]</td>
<td>![Speed Symbol]</td>
<td>![Speed Symbol]</td>
</tr>
</tbody>
</table>


#### Qualitative

Also called:  
- Categorical Attributes  
- Identity Channels

#### Quantitative

Also called:  
- Ordered Attributes  
- 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 SPOKE (Scalable Precision Medicine Oriented Knowledge Engine) 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 2033303.

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

Katy Börner1,2, Sarah A. Teichmann1, Ellen M. Quardokus3, James C. Gee4, Kristen Browne5, David Osumi-Sutherland6, Bruce W. Herr II7, Andreas Buell8, Brishkesh Paul9, Muzhifah Haniffa10, Laura Jardine, Amy Bernard1, Song-Lin Ding8, Jeremy A. Miller1, Shin Lin9, Marc K. Halushka11, Avinash Boppanna12, Teri A. Longacre13, John Hickey13, Yiing Lin13, M. Todd Valerius14, Yongquan He15, Gloria Pryhuber16, Xin Sun17, Marda Jorgensen18, Andrea J. Radtke19, Clive Waterfall18, Fiona Ginty20, Jonhan Ho21, Joel Sunshine22, Rebecca T. Beuschel22, Maigan Brusko21, Su Jin Lee23, Rajeev Malhotra14,13, Sanjay Jain24,25 and Griffin Weber26

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+3). 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 3-D 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 fund-
Specimen, Biological Structure, and Spatial Ontologies in Support of a Human Reference Atlas

https://biorxiv.org/cgi/content/short/2022.09.08.507220v1
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/