CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Registering, Visualizing, and Exploring Biomedical Data

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

NIMH Workshop on Advanced Statistical Methods and Dynamic Data Visualizations for Mental Health Studies

July 30, 2021

Overview

Mapping Science: An Exhibit

Mapping SPOKE: 3M Nodes and 30M Edges

HuBMAP: Toward a Human Reference Map

Data Visualization Literacy Framework

Empower Yourself!

CNS Cyberinfrastructure for Network Science Center

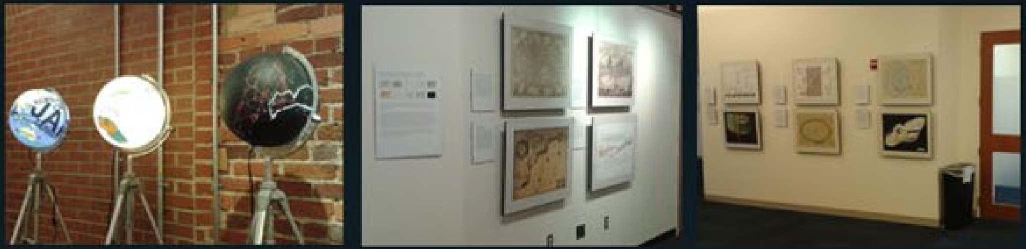
INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Mapping Science Exhibit

http://scimaps.org

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

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



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

http://scimaps.org

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

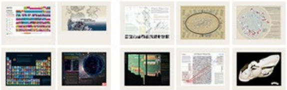
Places & Spaces: Mapping Science Exhibit

1st Decade (2005-2014)

Maps

Iteration I (2005)
The Power of Maps

Iteration II (2006)
The Power of Reference Systems



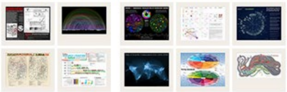
Iteration III (2007)
The Power of Forecasts

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 Kids

Iteration IX (2013)
Science Maps Showing Trends and Dynamics

Iteration X (2014)
The Future of Science Mapping

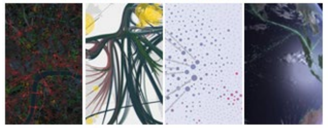
2nd Decade (2015-2024)

Macroscopes

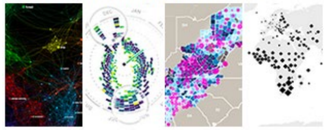
Iteration XI (2015)
Macroscopes for Interacting with Science

Iteration XIII (2017)
Macroscopes for Playing with Scale

Iteration XII (2016)
Macroscopes for Making Sense of Science



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



http://scimaps.org

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

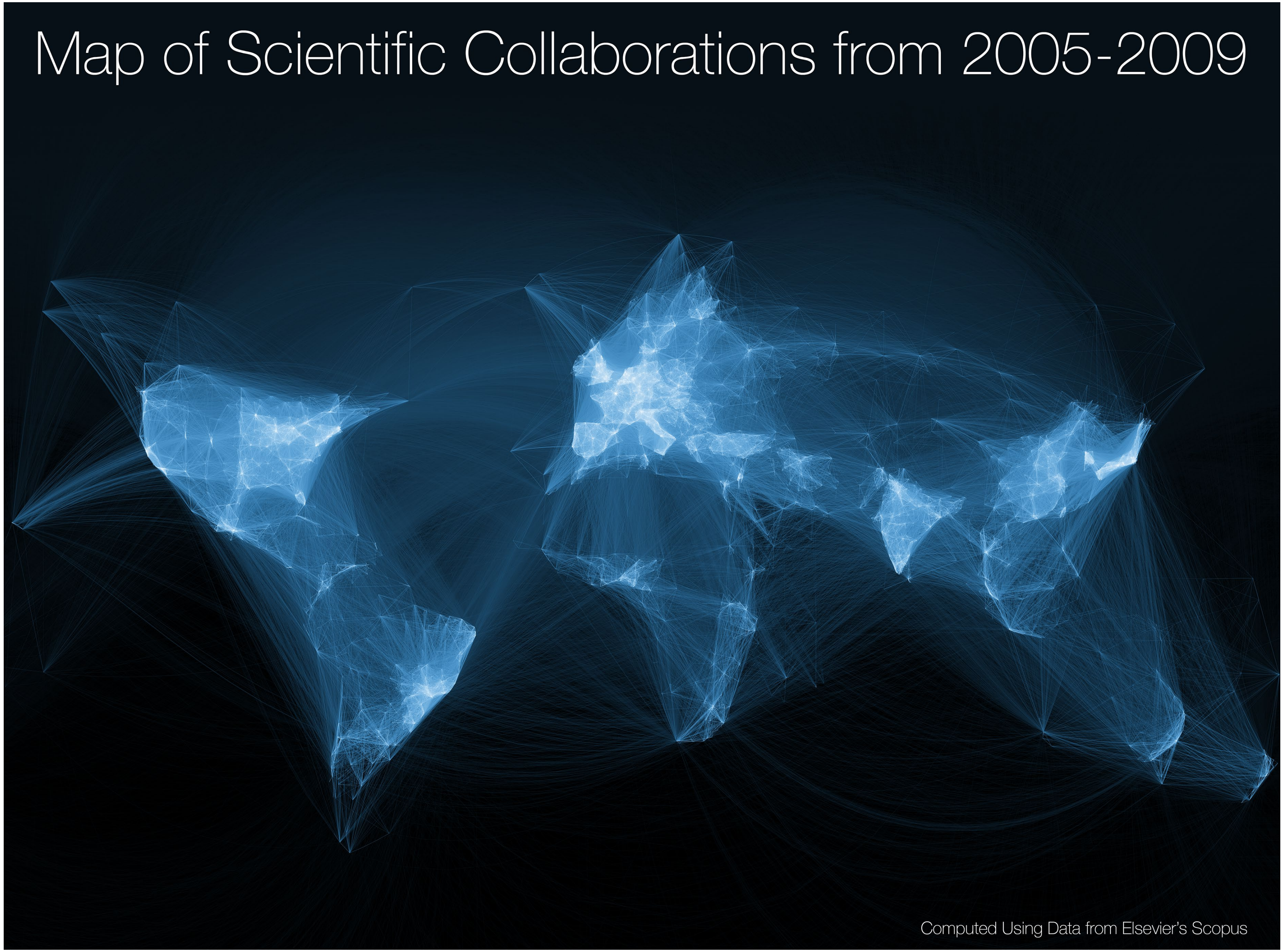
248
MAPMAKERS
from fields as disparate as art, urban planning, engineering, and the history of science.

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*.



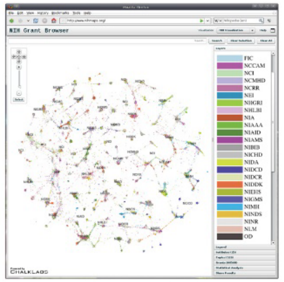
VII.6 Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012

A Topic Map of NIH Grants 2007

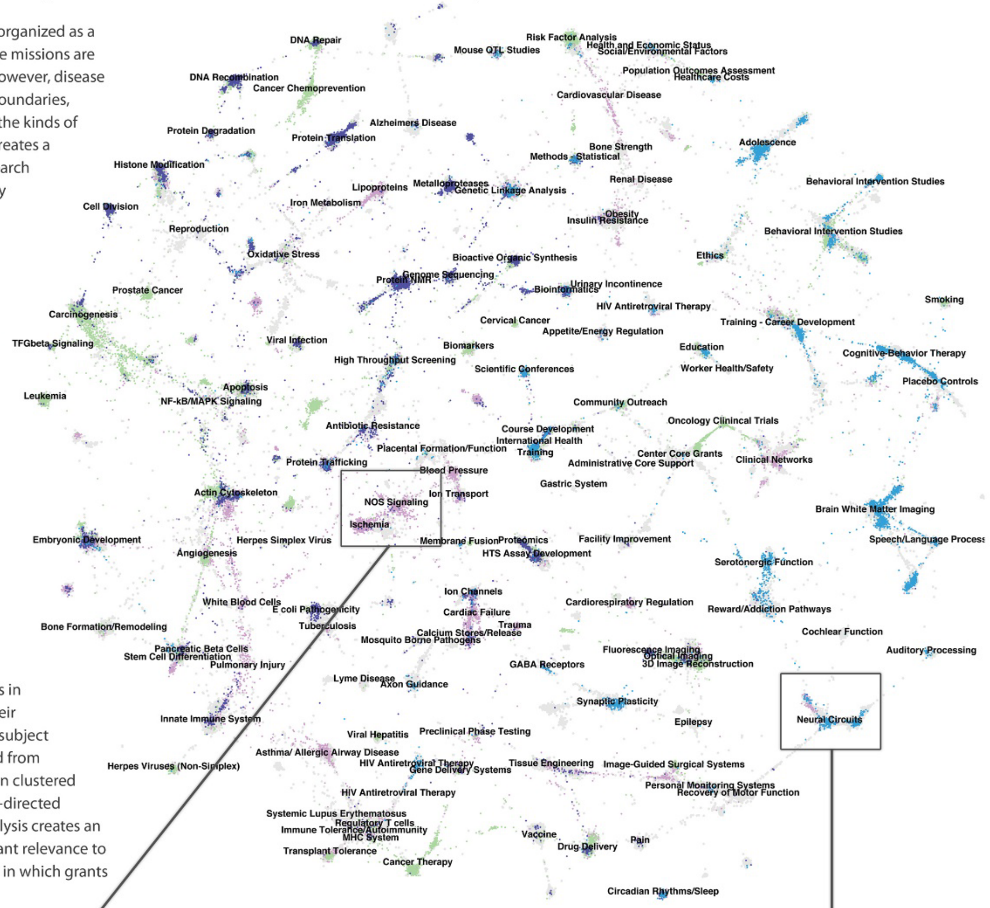
ChalkLabs ISI UCIRVINE

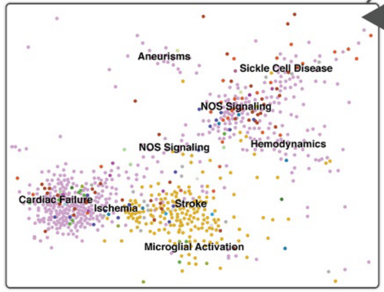
Bruce W. Herr II (Chalklabs & IU), Gully Burns (ISI), David Newman (UCI), Edmund Talley (NIH)

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



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



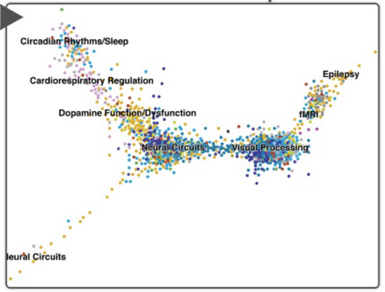


Cardiac Diseases Research

An area of the map focused on cardio-vascular function and dysfunction. Cardiac Failure (primarily funded by NHLBI) is topically clustered next to Stroke (NINDS), since these are the two major medical emergencies associated with ischemia, which results from a re-stricted blood supply. Also localized in this area are grants focused on Nitric Oxide (NOS) Signaling, a major biochem-ical pathway for vasodilation, and grants on Hemodynamics, Sickle Cell Disease, and Aneurysms.

Neural Circuits Research

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



National Cancer Institute (NCI)

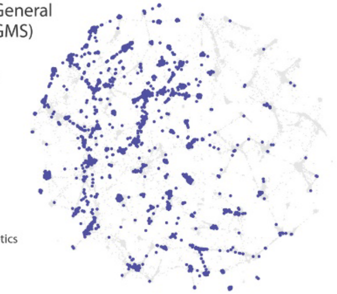
TOP 10 TOPICS

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

National Institute of General Medical Sciences (NIGMS)

TOP 10 TOPICS

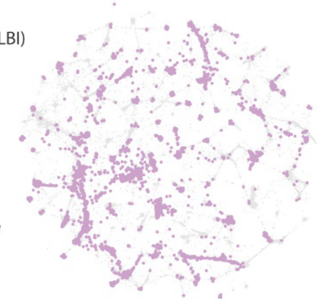
1. Bioactive Organic Synthesis
2. X-ray Crystallography
3. Protein NMR
4. Computational Models
5. Yeast Biology
6. Metalloproteases
7. Enzymatic Mechanisms
8. Protein Complexes
9. Invertebrate/Zebrafish Genetics
10. Cell Division



National Heart, Lung, and Blood Institute (NHLBI)

TOP 10 TOPICS

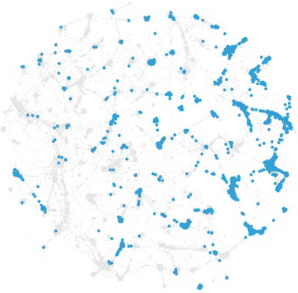
1. Cardiac Failure
2. Pulmonary Injury
3. Genetic Linkage Analysis
4. Cardiovascular Disease
5. Atherosclerosis
6. Hemostasis
7. Blood Pressure
8. Asthma/ Allergic Airway Disease
9. Gene Association
10. Lipoproteins



National Institute of Mental Health (NIMH)

TOP 10 TOPICS

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



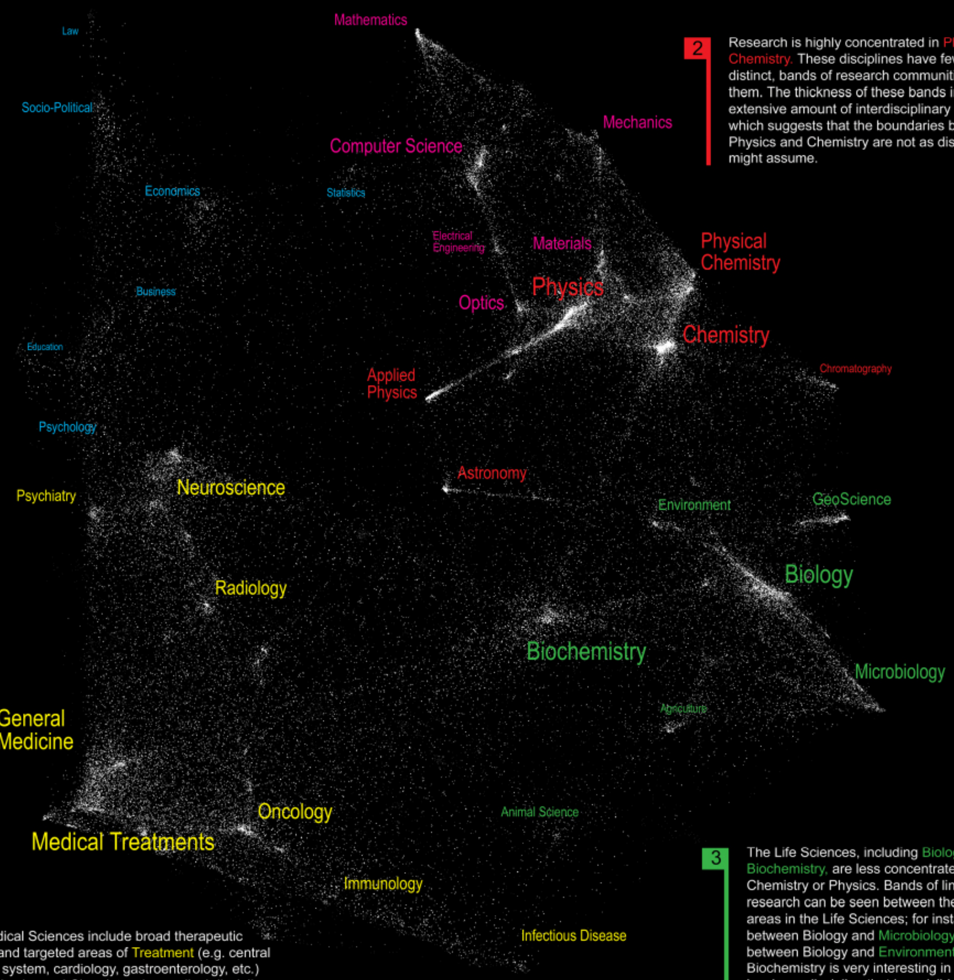
V.7 *A Topic Map of NIH Grants 2007* - Bruce W. Herr II, Gully A.P.C. Burns, David Newman, and Edmund Talley - 2009

The Structure of Science

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

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

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



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

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

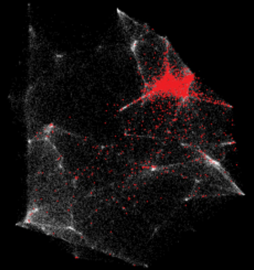
We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

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

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

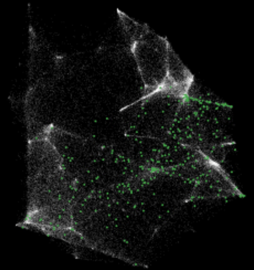
The map of science can be used as a tool for science strategy. This is the terrain in which organizations and institutions locate their scientific capabilities. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to explore,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

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



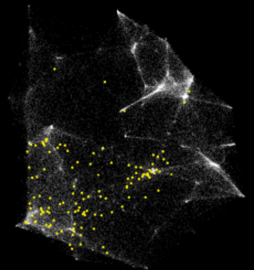
Nanotechnology

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



Proteomics

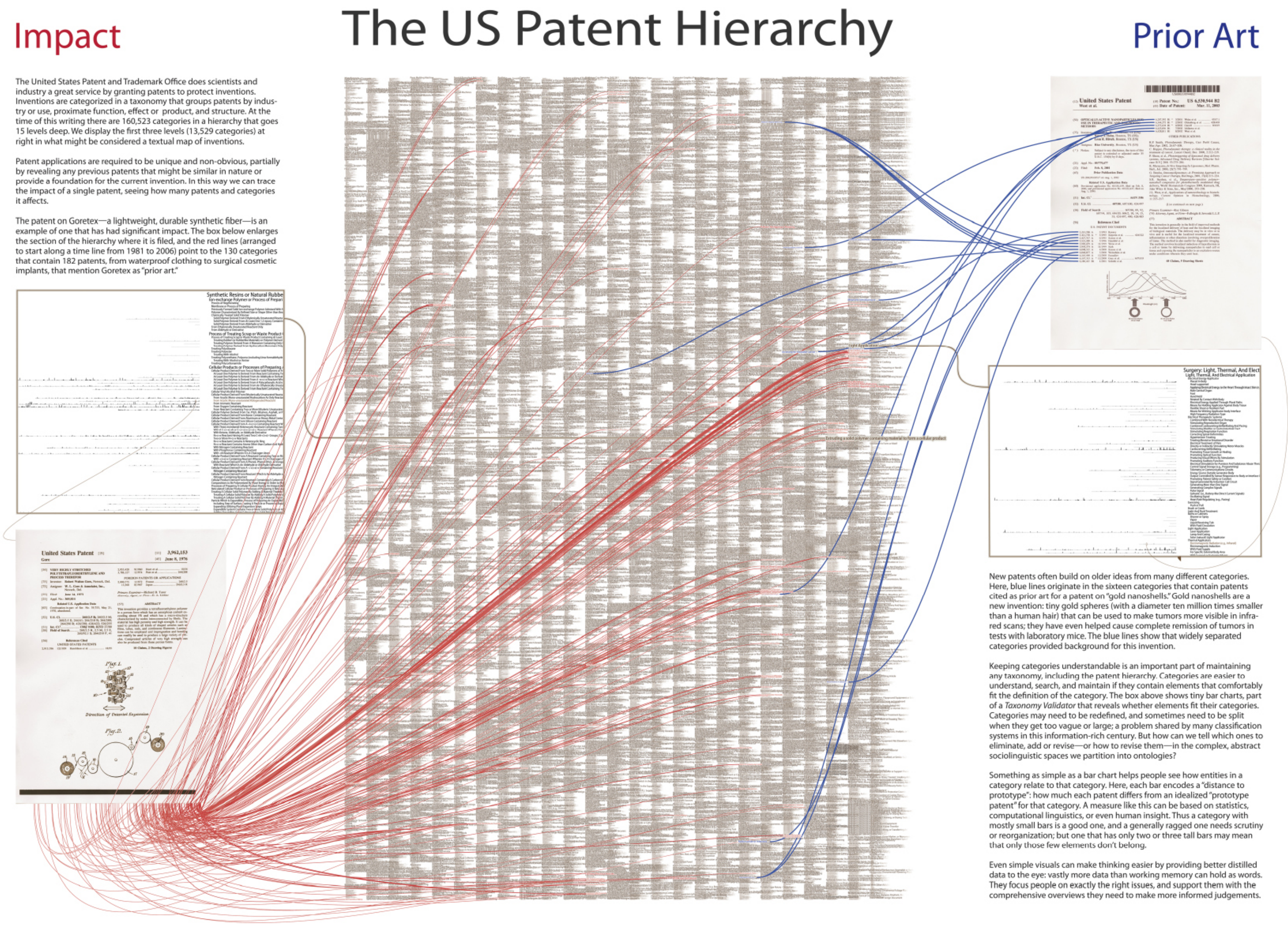
Research communities in proteomics are centered in Biochemistry. In addition, there is a heavy focus in the tools section of chemistry, such as Chromatography. The balance of the proteomics communities are widely dispersed among the Life and Medical Sciences.



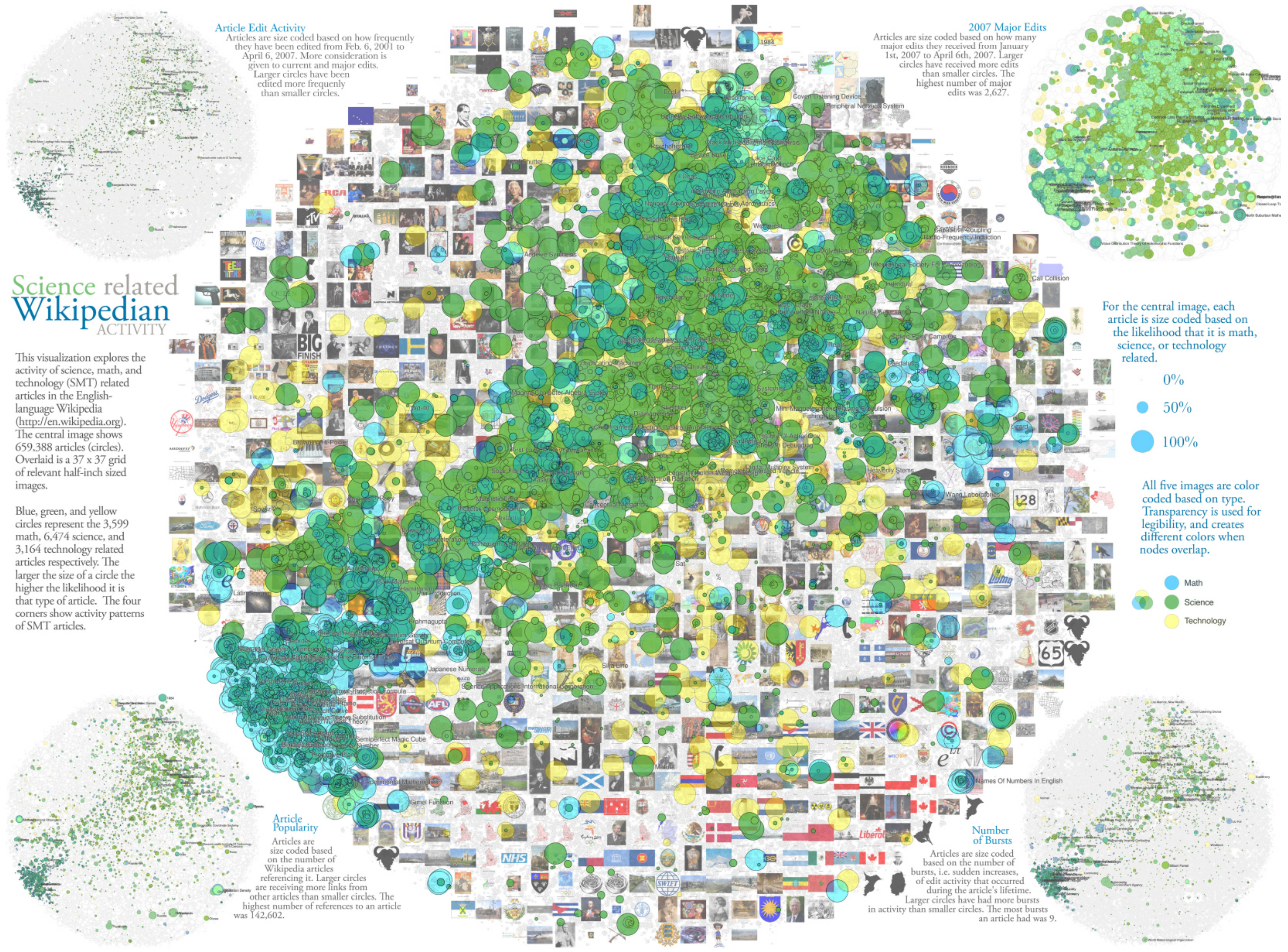
Pharmacogenomics

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

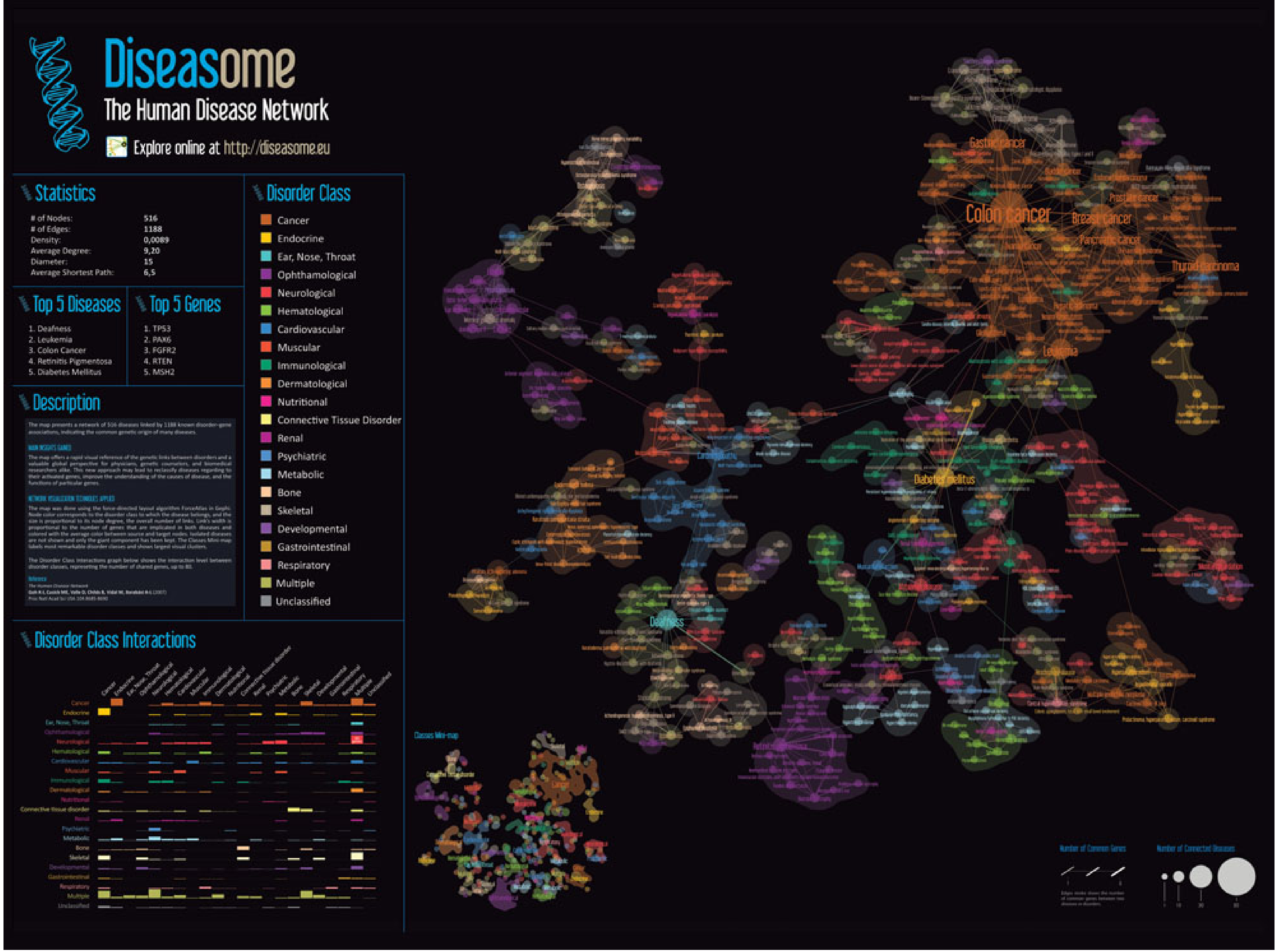
I.10 *The Structure of Science* - Kevin W. Boyack and Richard Klavans - 2005



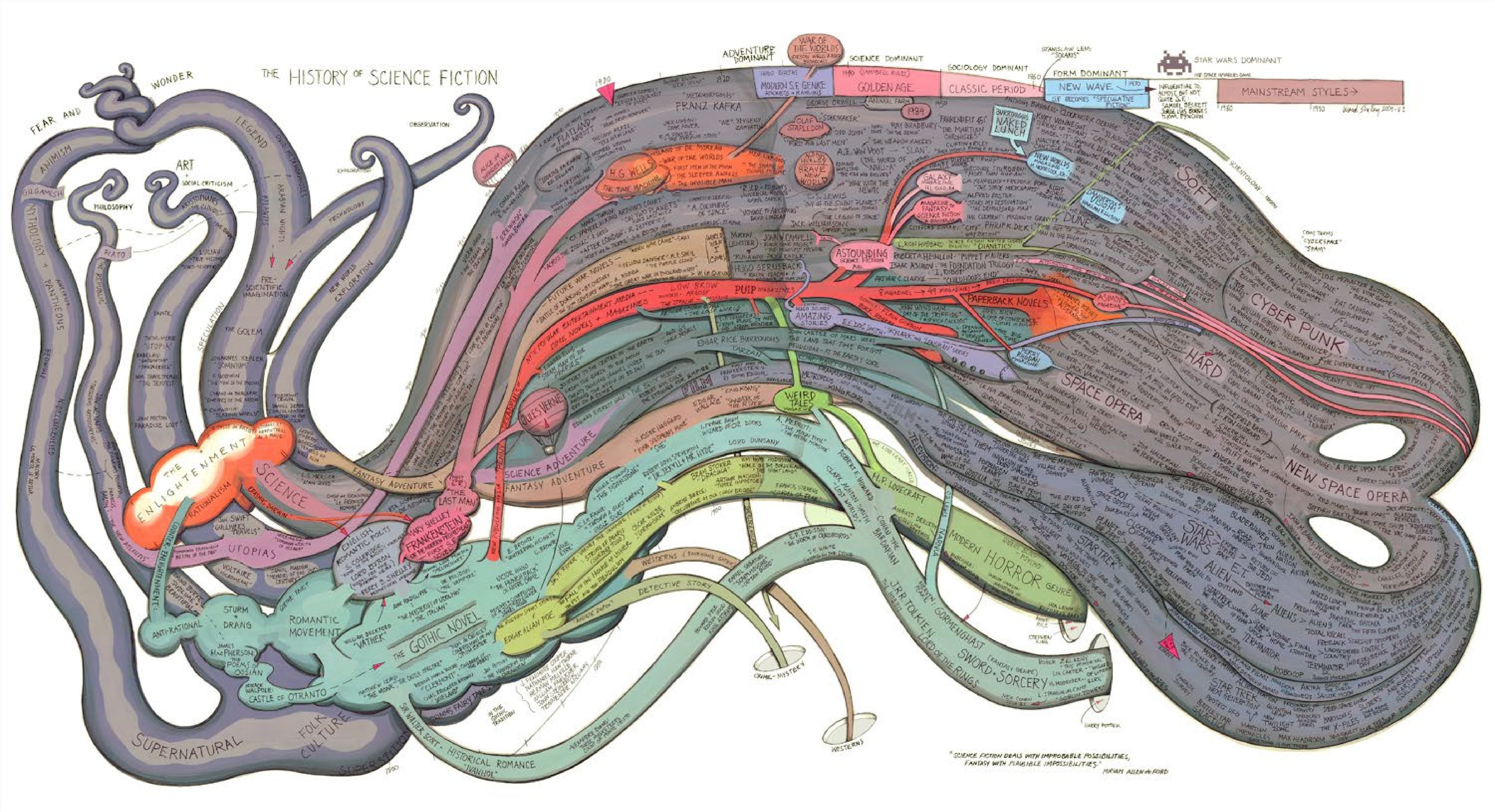
II.8 *Taxonomy Visualization of Patent Data* - Katy Börner, Elisha F. Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley - 2006



III.8 *Science-Related Wikipedian Activity* - Bruce W. Herr II, Todd M. Holloway, Elisha F. Hardy, Katy Börner, and Kevin Boyack - 2007

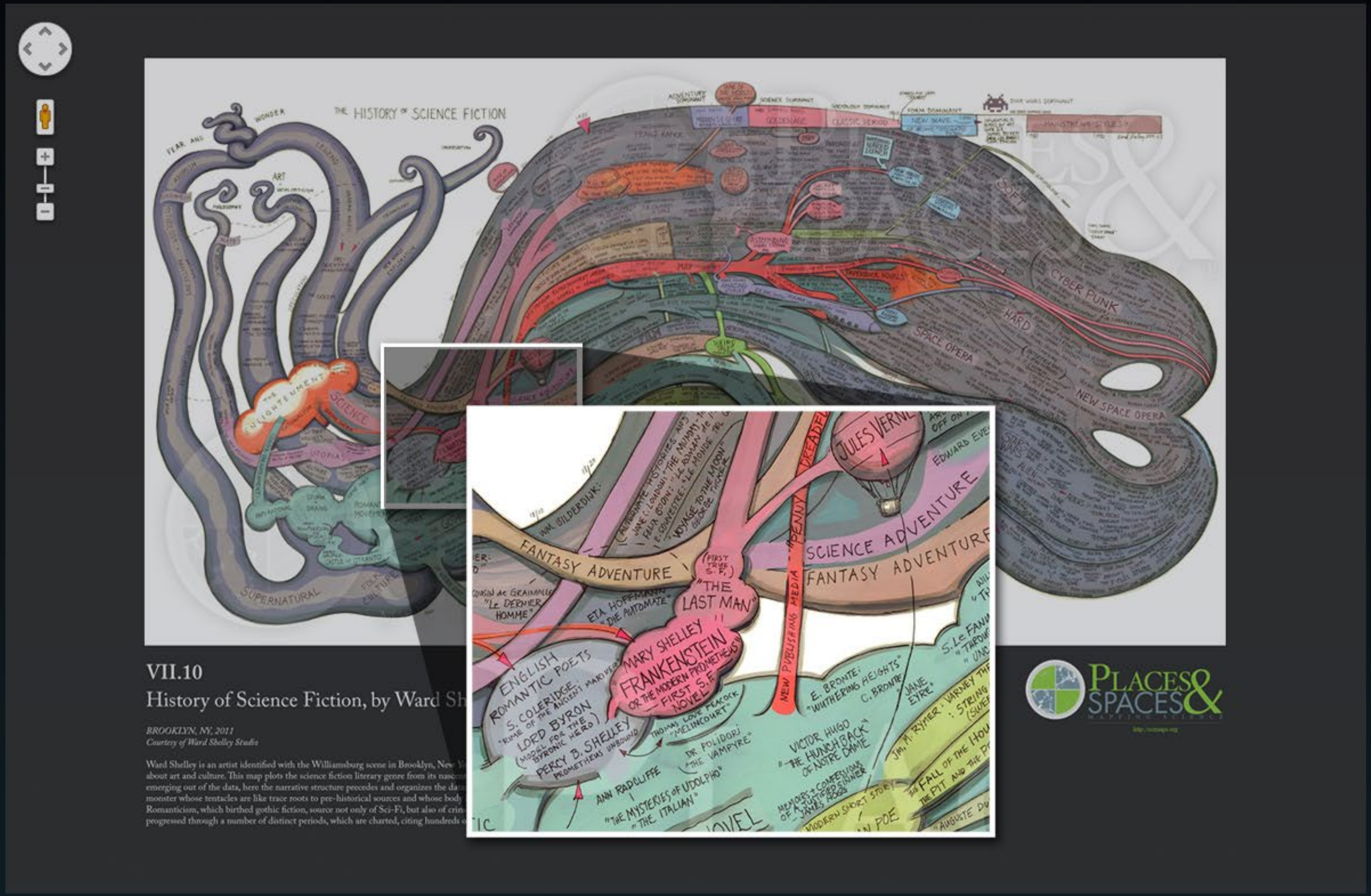


VI.3 *Diseasome: The Human Disease Network* - Mathieu Bastian and Sébastien Heymann - 2009



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

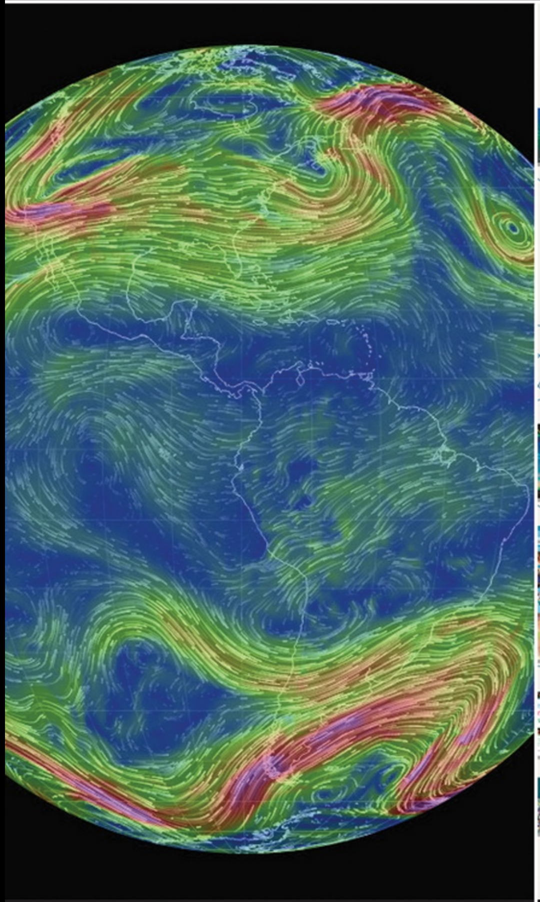
Check out our Zoom Maps online!



Visit scimaps.org and check out all our maps in stunning detail!

MACROSCOPES FOR INTERACTING WITH SCIENCE

PLACES & SPACES
MAPPING SCIENCE
scimaps.org

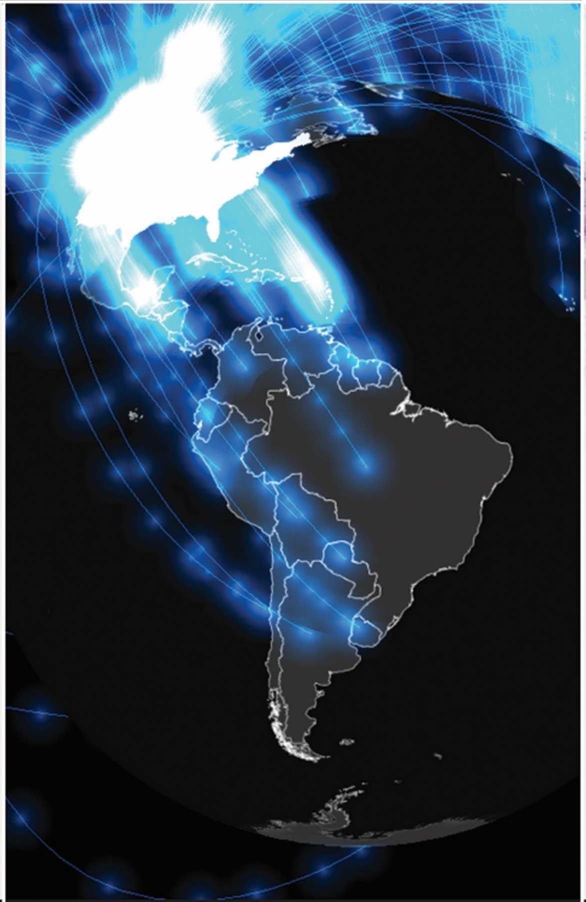


Earth

Weather on a worldwide scale

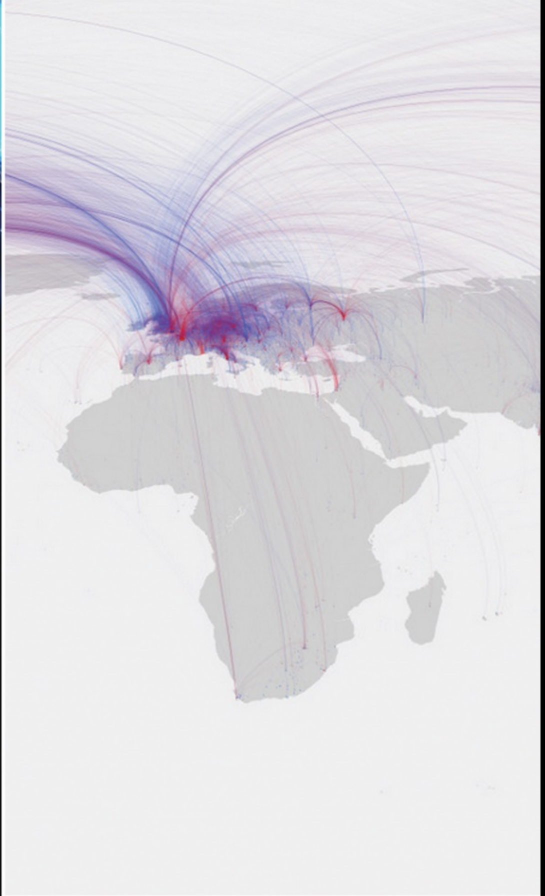
AcademyScope

Exploring the scientific landscape



Mapping Global Society

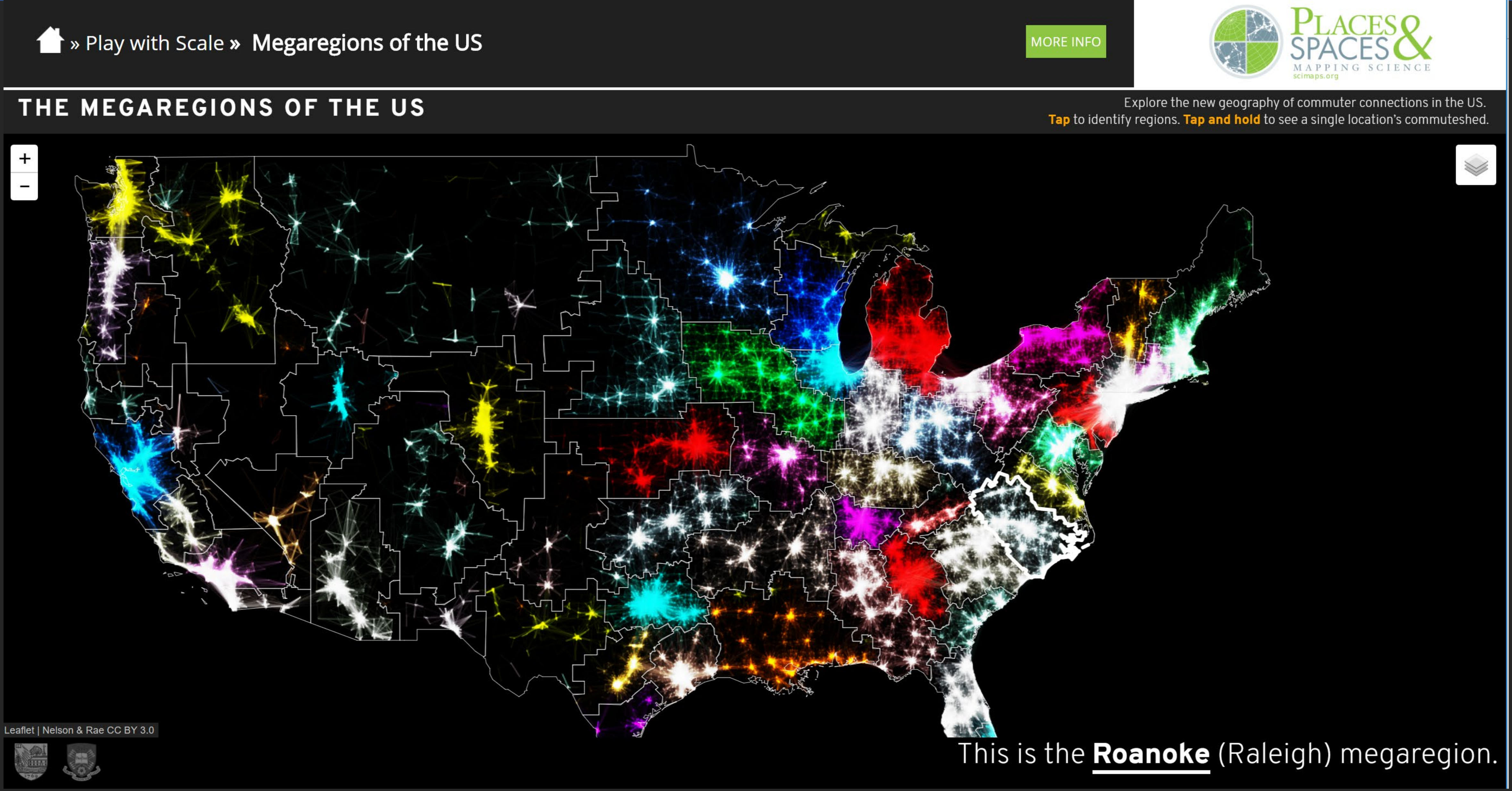
Local news from a global perspective



Charting Culture

2,600 years of human history in 5 minutes

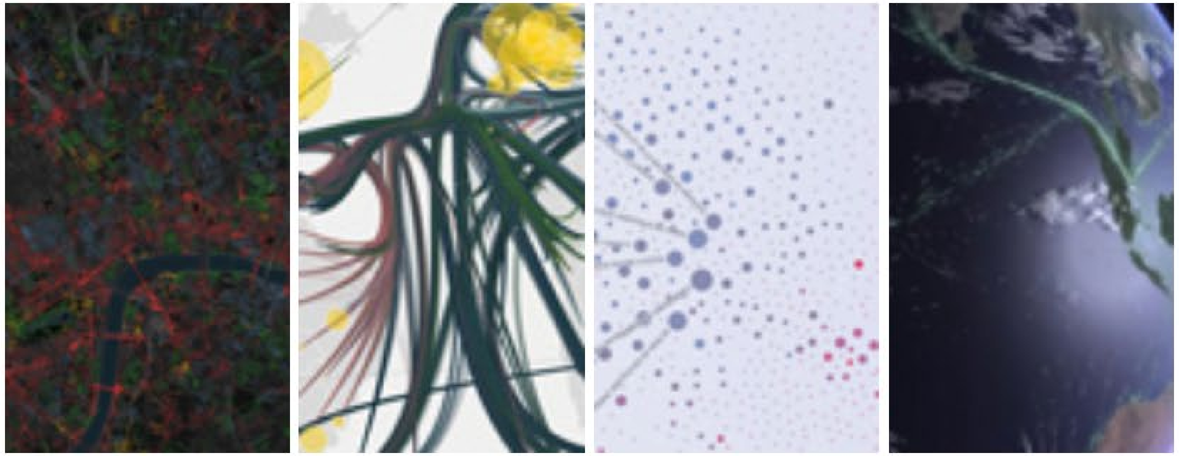
Smelly Maps – Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello – 2015



Megaregions of the US –Garrett Dash Nelson and Alasdair Rae – 2016

Iteration XII (2016)

Macroscopes for Making Sense of Science

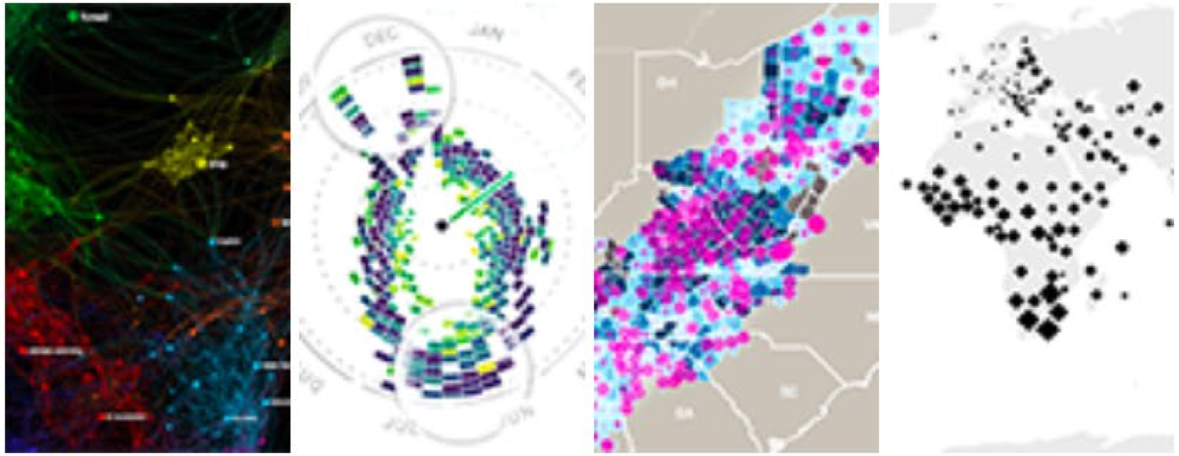


Iteration XIII (2017)

Macroscopes for Playing with Scale

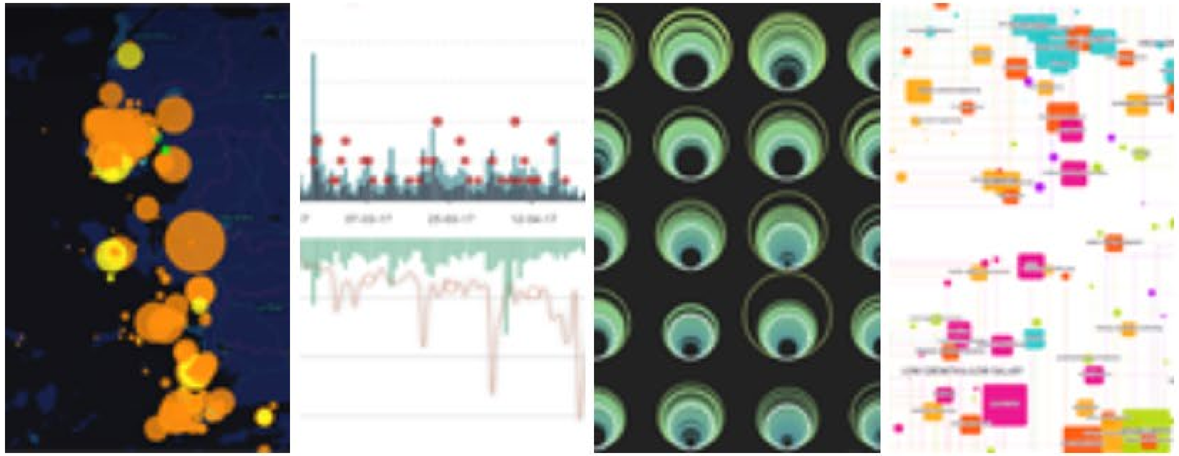
Iteration XIV (2018)

Macroscopes for Ensuring our Well-being



Iteration XV (2019)

Macroscopes for Tracking the Flow of Resources



Acknowledgments

Exhibit Curators

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

http://scimaps.org

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

Exhibit Advisory Board

Gary Berg-Cross

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

Donna J. Cox, MFA, Ph.D.

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

Bonnie DeVarco

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

Peter A. Hook

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

Francis Harvey

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

Lev Manovich

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

André Skupin

Associate Professor of Geography at San Diego State University, California

Moritz Stefaner

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

Olga Subirós

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

Stephen Uzzo

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

Benjamin Wiederkehr

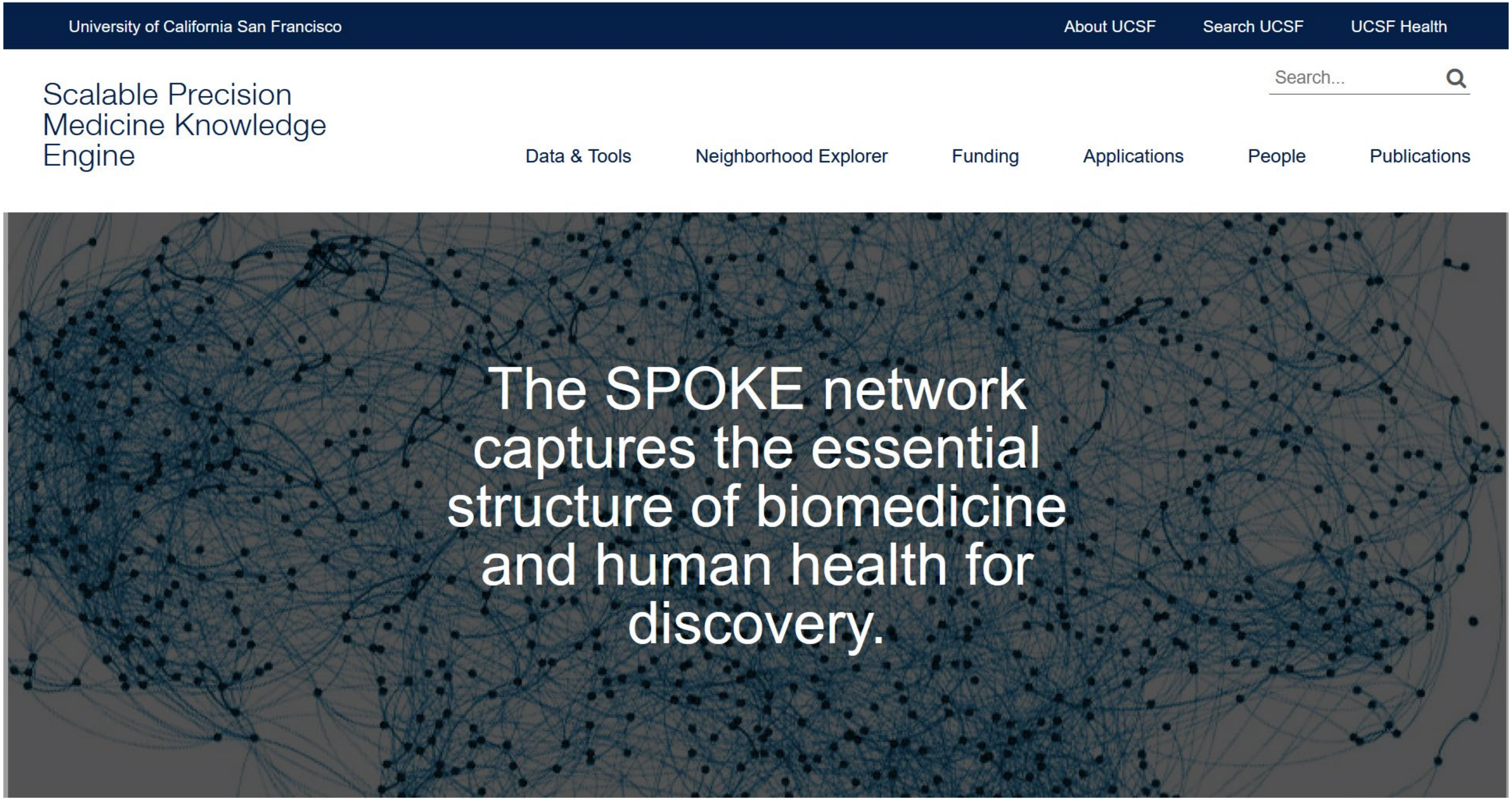
Founding Partner and Managing Director of Interactive Things in Zürich, Switzerland

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Visualizations of the Scalable Precision Medicine Knowledge Engine (SPOKE)

https://spoke.ucsf.edu



https://spoke.ucsf.edu

Lead Investigators

Sergio Baranzini, PhD
Principal Investigator

Sui Huang, MD, PhD (ISB)

Sharat Israni, PhD

Mike Keiser, PhD

Technical & Planning Team

Rafael Gonçalves, PhD (Stanford)

Adil Harroud, MD

Elaine Meng, MD

Scooter Morris, PhD

Charlotte Nelson, PhD

Boris Oskotsky, PhD

Angela Rizk-Jackson, PhD

Peter Rose, PhD (UCSD)

Brett Smith (ISB)

Karthik Soman, PhD

Xiaoyuan Zhou, PhD

Collaborators

Katy Börner, PhD (IU)

William Brown, PhD, DrPH

Ramanathan V. Guha, PhD (Google)

Mark Musen, MD, PhD (Stanford)

Camille Nebeker, EdD, MS (UCSD)

Roger Pearce, PhD (LNL)

SPOKE investigative teams

The SPOKE team members are from the following organizations. *Team members listed below are from UCSF, except when indicated.*

- Google
- Indiana University (IU)
- Institute for Systems Biology (ISB)
- Lawrence Livermore National Lab (LLNL)
- Stanford University
- University of California, San Diego (UCSD)
- University of California, San Francisco (UCSF)

SPOKE Scalable Precision Medicine Knowledge Engine

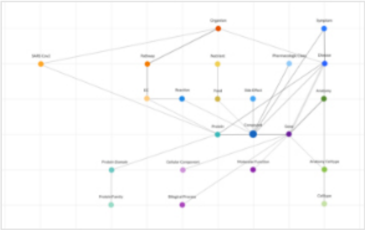
Envisioning SPOKE: 3M Nodes and 30M Edges

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.

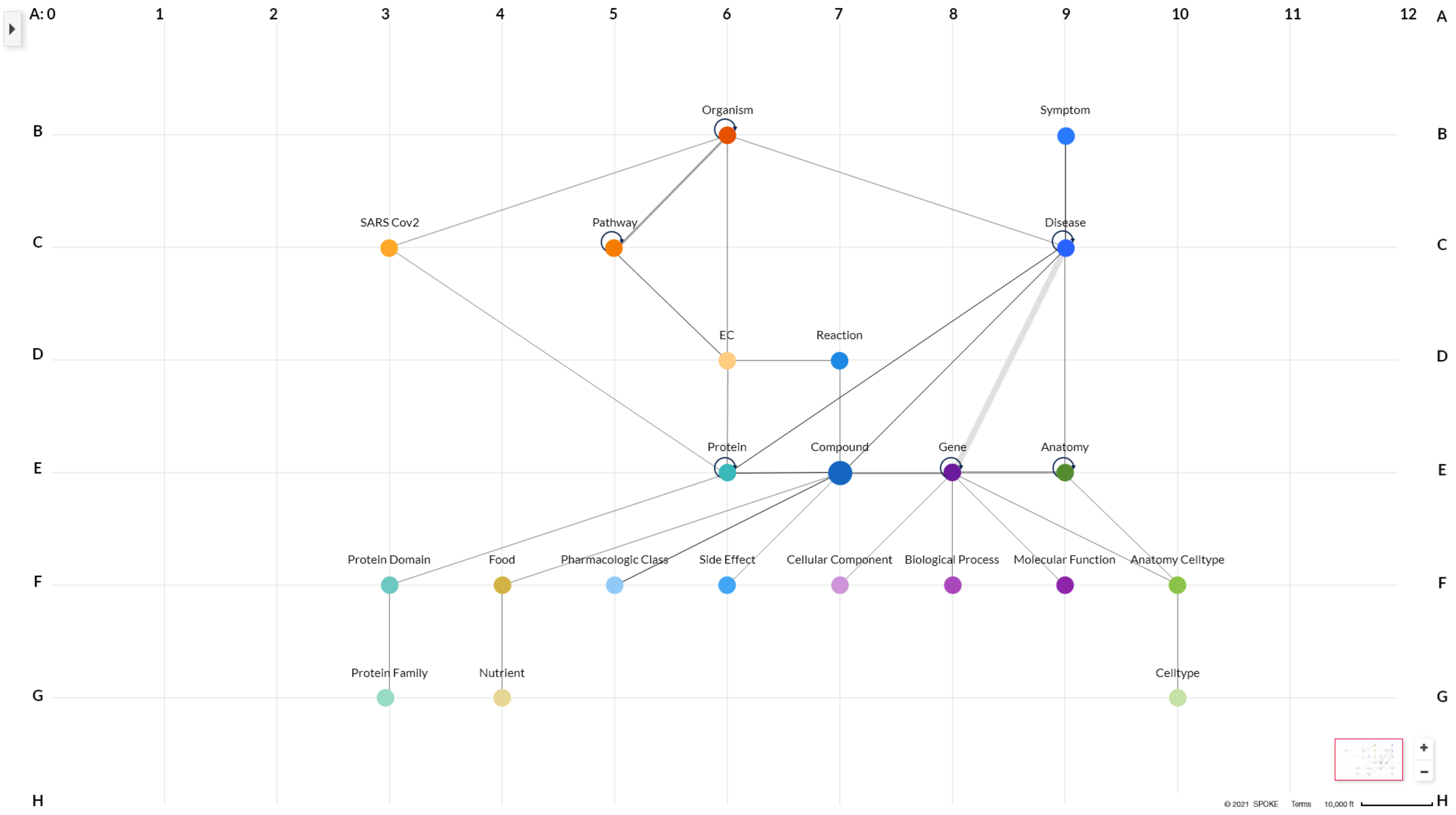
The site was designed for two user groups: (1) novice users interested to understand the coverage and quality of SPOKE data and (2) expert users interested to analyze and optimize the interlinked knowledge graphs in SPOKE.
The overview visualization shows the different entity type and their diverse interlinkages. Detail

SPOKE is a fully interactive tool for exploring the interconnections between data.

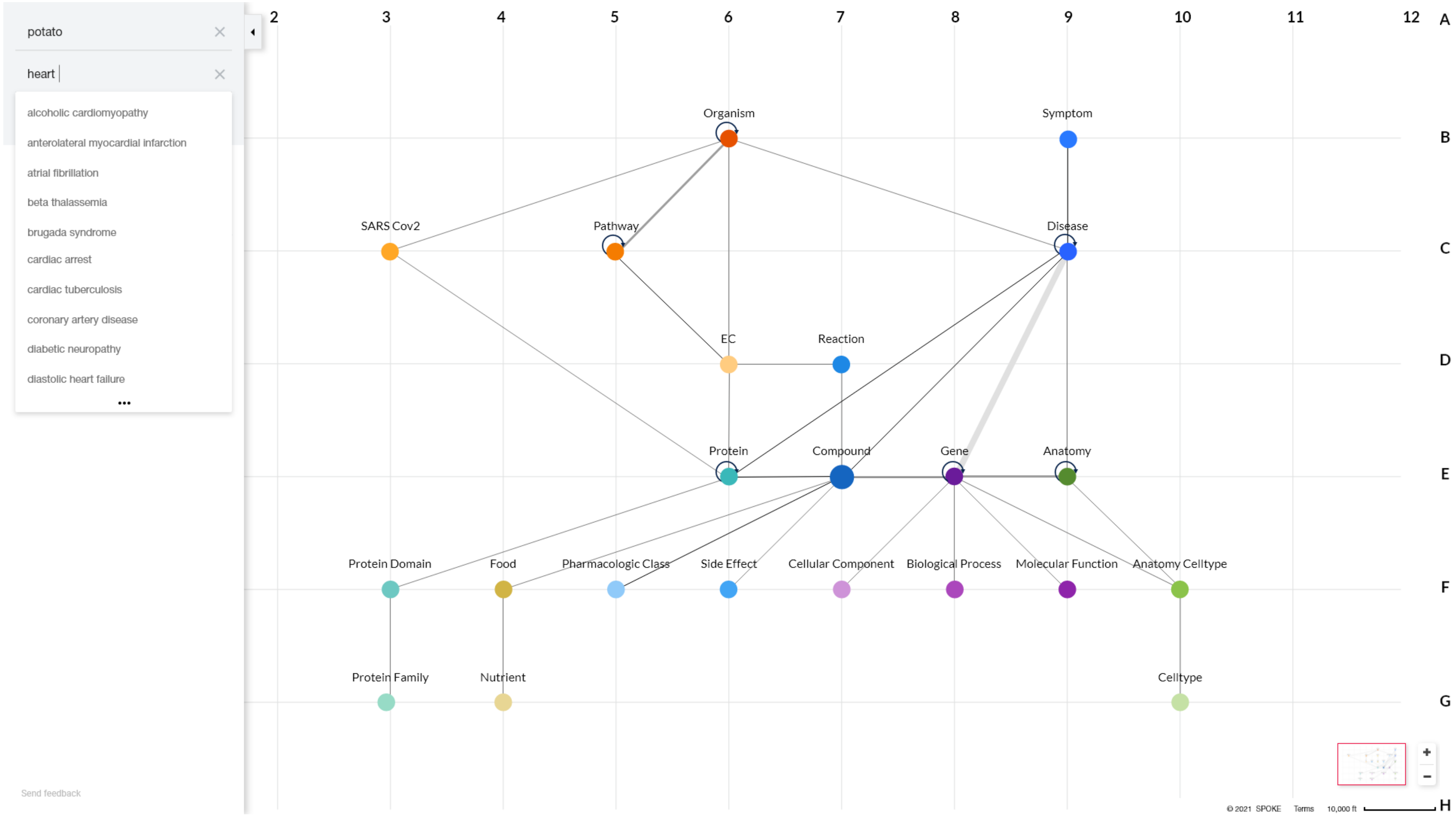
Explore SPOKE

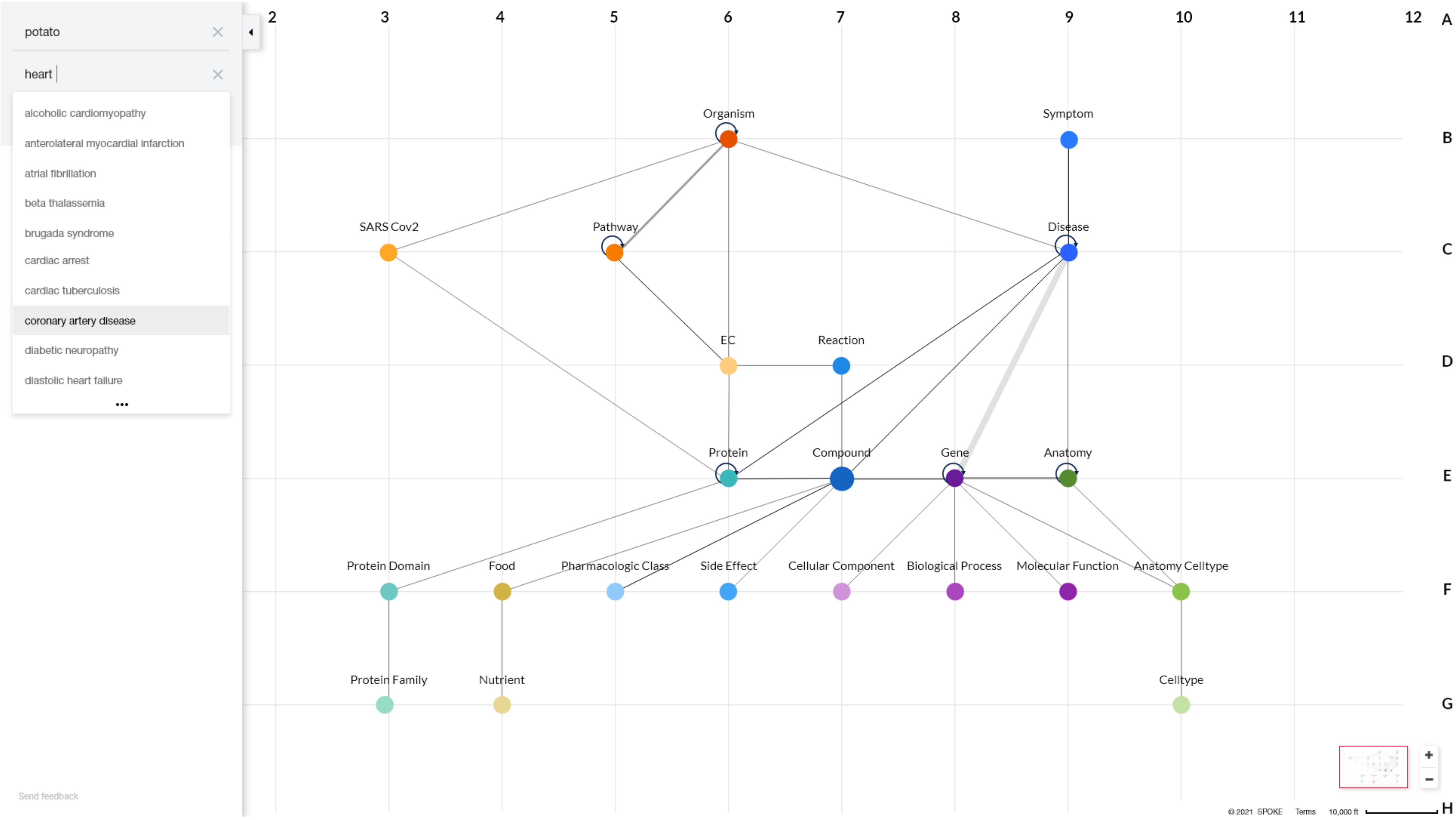


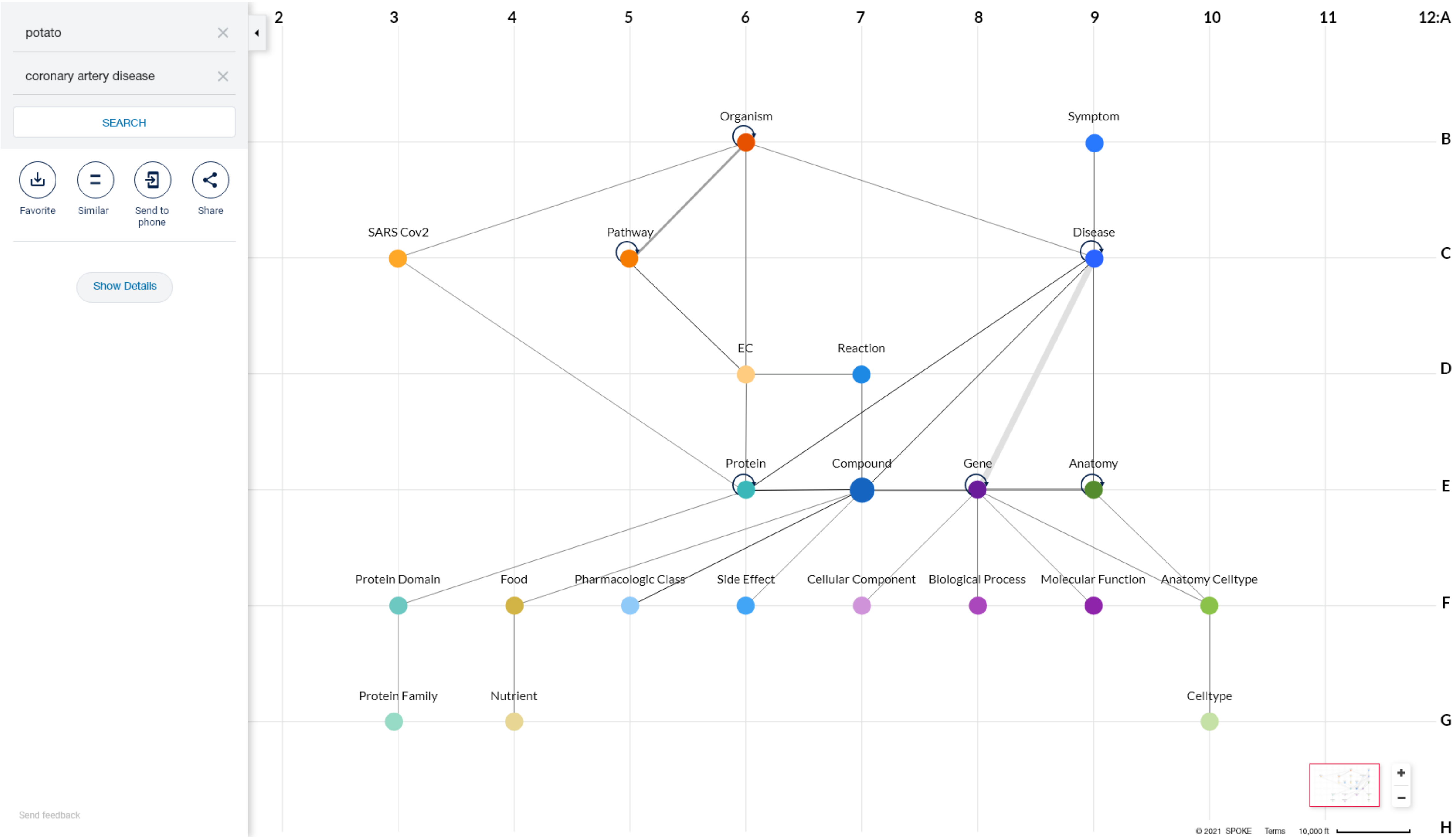
NSF NIH National Center for Advancing Translational Sciences UCSF CNS Cyberinfrastructure for Network Science Center

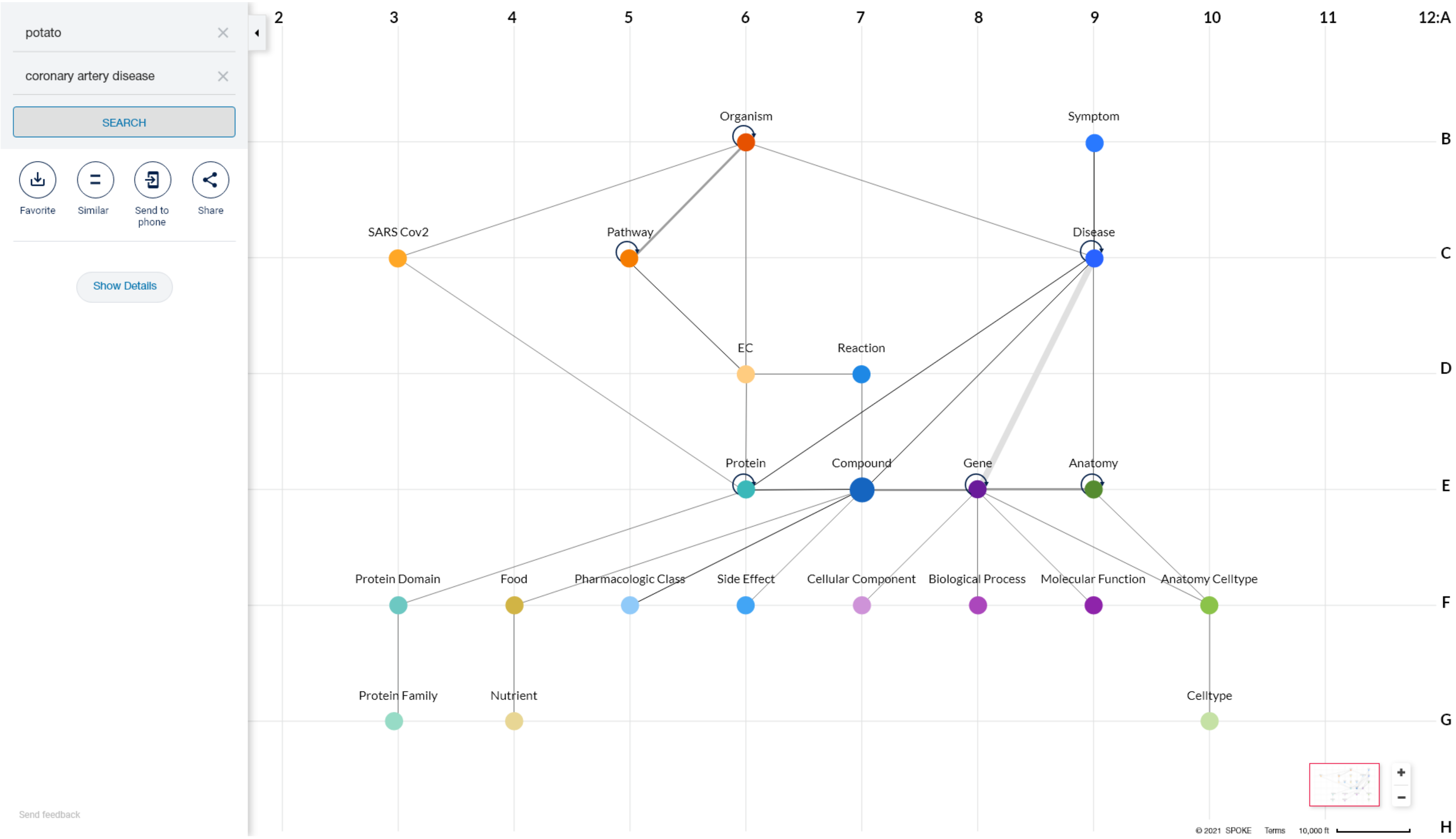


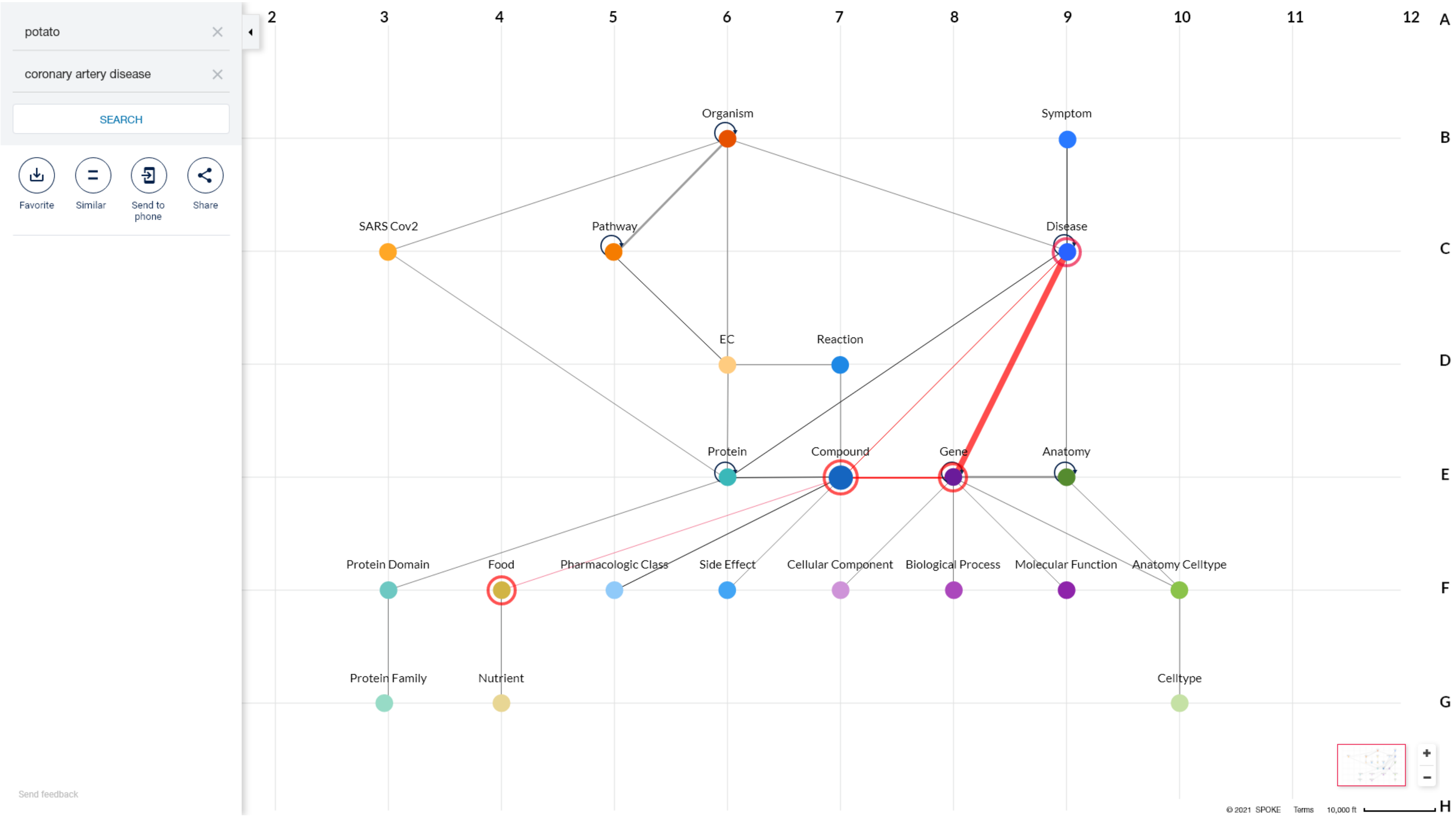
A: 0
1
2
3
4
5
6
7
8
9
10
11
12
A
B
C
D
E
F
G
H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft

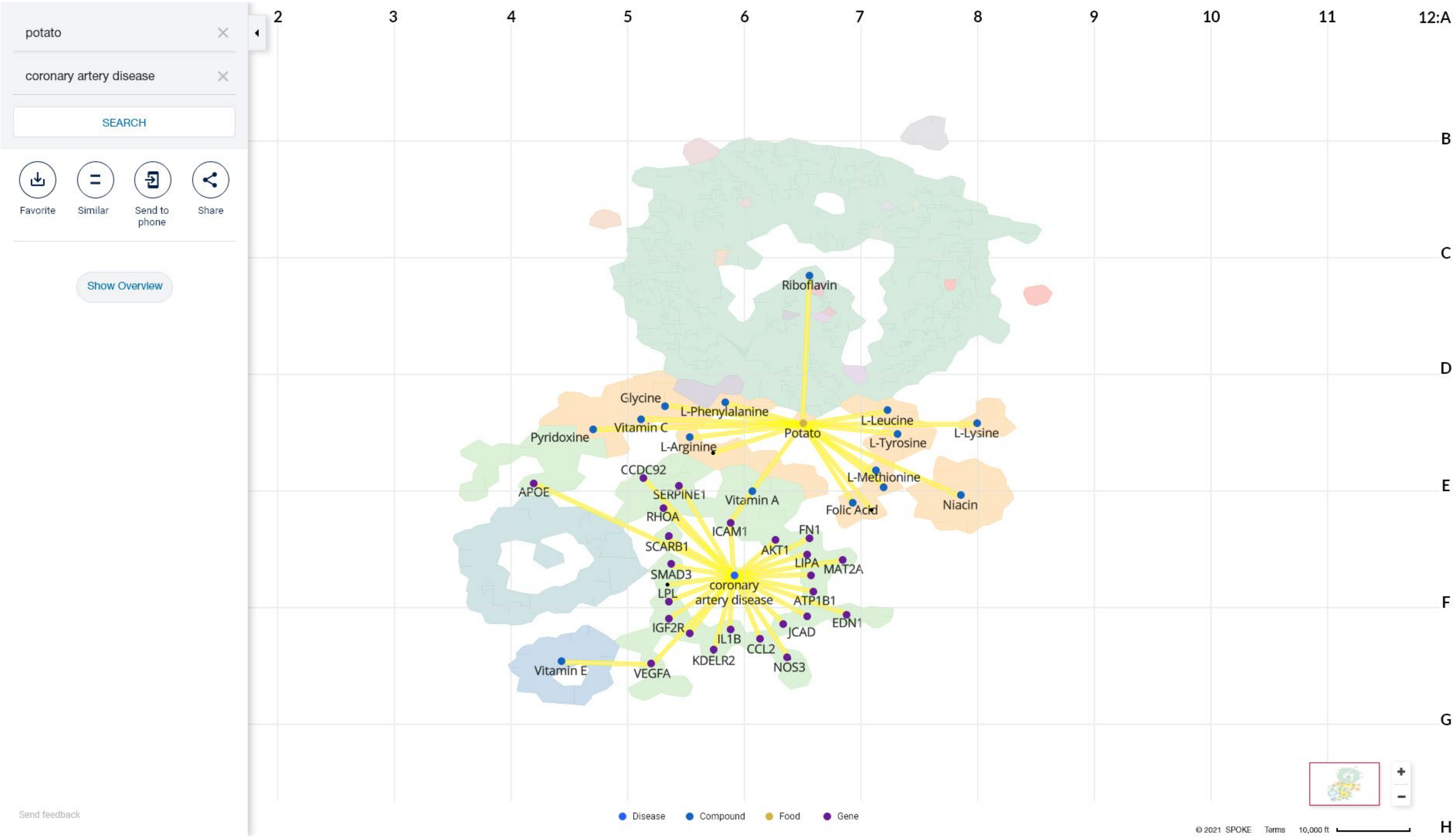

potato
heart
alcoholic cardiomyopathy
anterolateral myocardial infarction
atrial fibrillation
beta thalassemia
brugada syndrome
cardiac arrest
cardiac tuberculosis
coronary artery disease
diabetic neuropathy
diastolic heart failure
...
Send feedback
2 3 4 5 6 7 8 9 10 11 12
A B C D E F G H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft

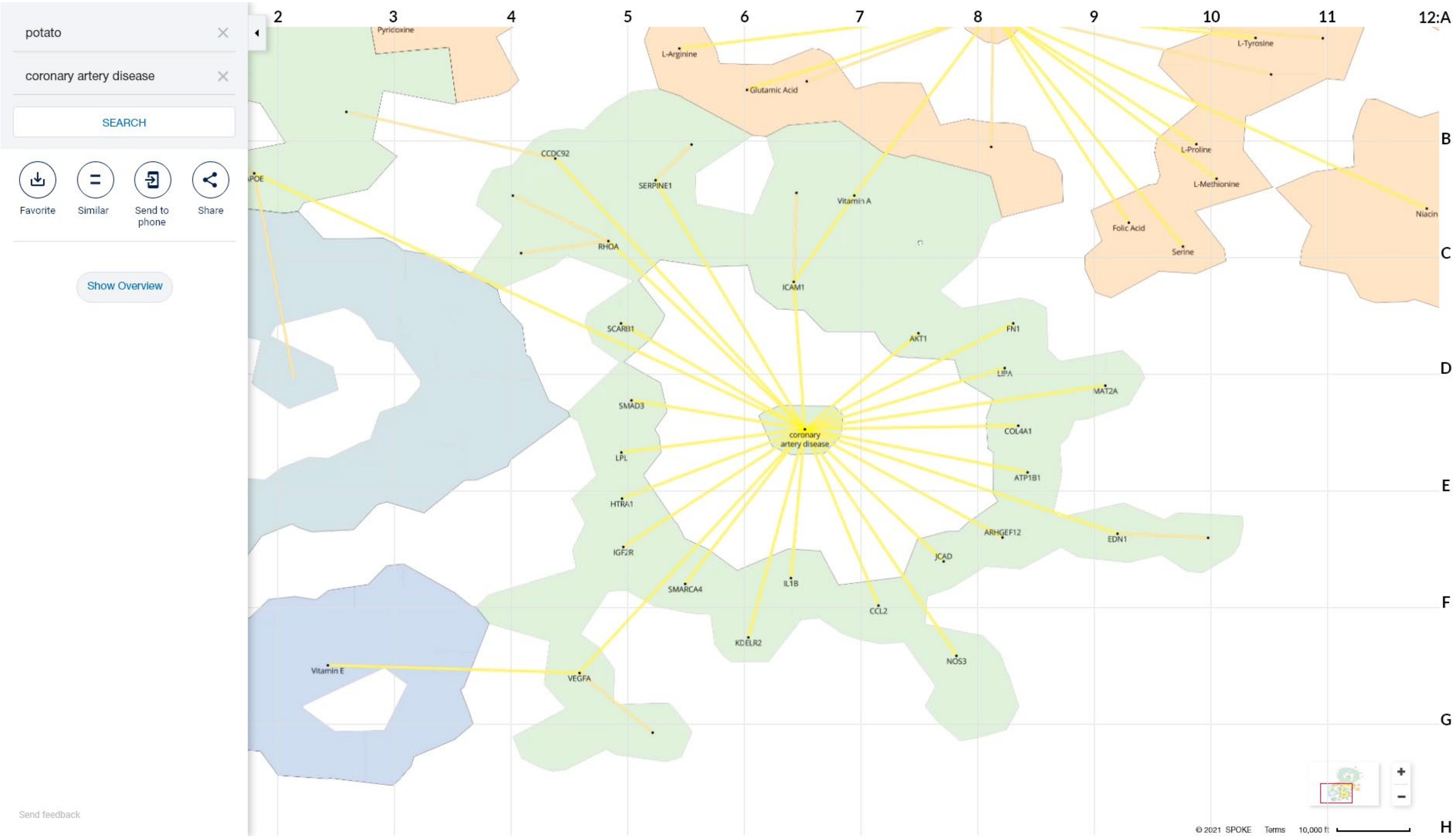

potato
heart
alcoholic cardiomyopathy
anterolateral myocardial infarction
atrial fibrillation
beta thalassemia
brugada syndrome
cardiac arrest
cardiac tuberculosis
coronary artery disease
diabetic neuropathy
diastolic heart failure
...
Send feedback
2
3
4
5
6
7
8
9
10
11
12
A
B
C
D
E
F
G
H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft

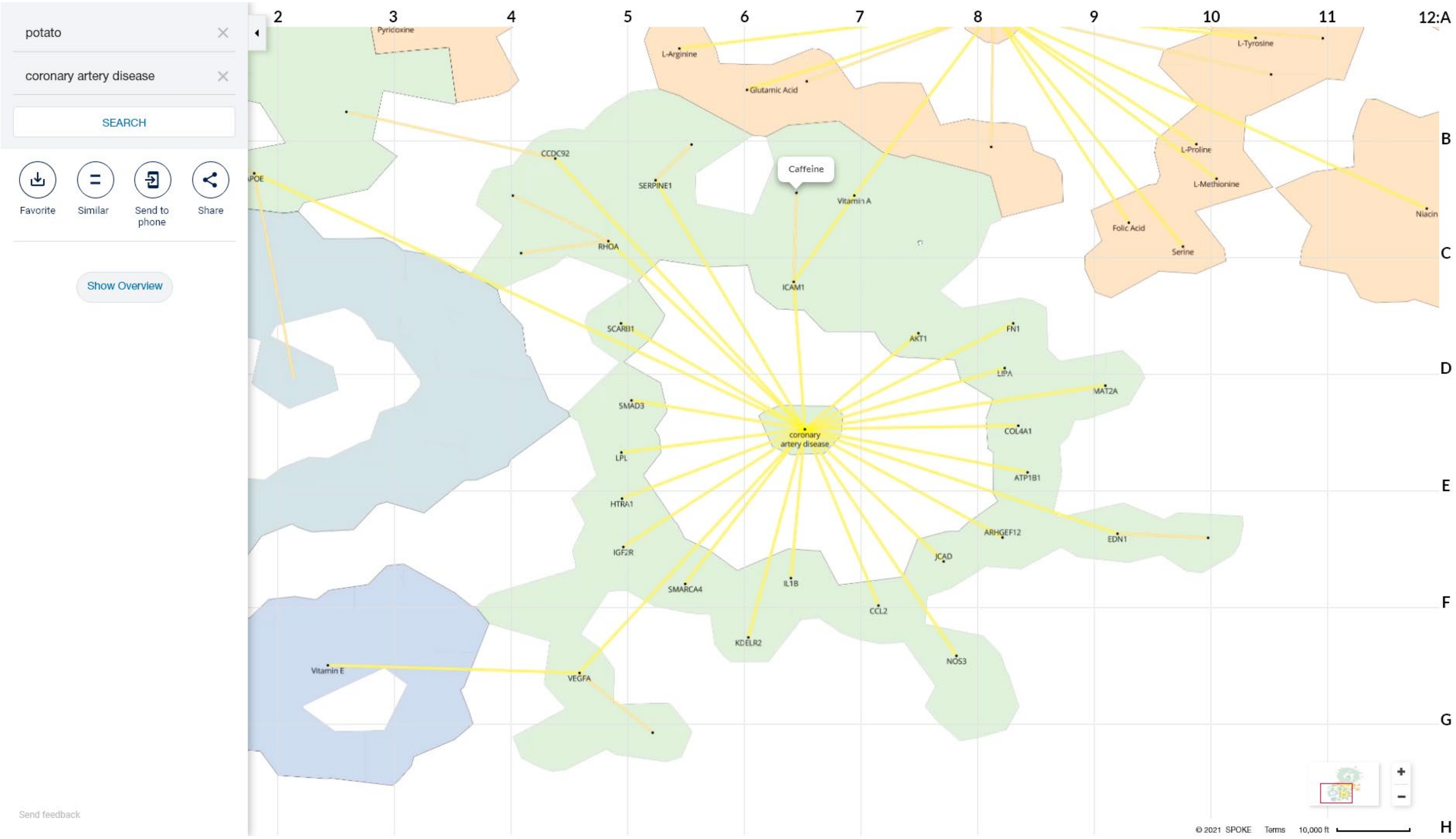

potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Show Details
Send feedback
2
3
4
5
6
7
8
9
10
11
12:A
B
C
D
E
F
G
H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft


potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Show Details
Send feedback
2
3
4
5
6
7
8
9
10
11
12:A
B
C
D
E
F
G
H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft


potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Send feedback
2
3
4
5
6
7
8
9
10
11
12
A
B
C
D
E
F
G
H
Organism
Symptom
SARS Cov2
Pathway
Disease
EC
Reaction
Protein
Compound
Gene
Anatomy
Protein Domain
Food
Pharmacologic Class
Side Effect
Cellular Component
Biological Process
Molecular Function
Anatomy Celltype
Protein Family
Nutrient
Celltype
© 2021 SPOKE
Terms
10,000 ft


potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Show Overview
Send feedback
2
3
4
5
6
7
8
9
10
11
12:A
B
C
D
E
F
G
H
Riboflavin
Glycine
L-Phenylalanine
Vitamin C
Pyridoxine
L-Arginine
Potato
L-Leucine
L-Lysine
L-Tyrosine
L-Methionine
Niacin
Folic Acid
Vitamin A
CCDC92
APOE
SERPINE1
RHOA
ICAM1
FN1
AKT1
SCARB1
LIPA
MAT2A
SMAD3
LPL
coronary artery disease
ATP1B1
EDN1
IGF2R
IL1B
JCAD
CCL2
KDELR2
NOS3
Vitamin E
VEGFA
Disease
Compound
Food
Gene
© 2021 SPOKE
Terms
10,000 ft


potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Show Overview
Send feedback
2
3
4
5
6
7
8
9
10
11
12:A
B
C
D
E
F
G
H
Pyridoxine
L-Arginine
Glutamic Acid
L-Tyrosine
L-Proline
L-Methionine
Niacin
Folic Acid
Serine
CCDC92
SERPINE1
Vitamin A
APOE
RHOA
ICAM1
SCARB1
AKT1
FN1
LIPA
MAT2A
SMAD3
coronary artery disease
COL4A1
LPL
ATP1B1
HTRA1
ARHGEF12
EDN1
IGF2R
JCAD
SMARCA4
IL1B
CCL2
KDELR2
NOS3
Vitamin E
VEGFA
© 2021 SPOKE
Terms
10,000 ft


potato
coronary artery disease
SEARCH
Favorite
Similar
Send to phone
Share
Show Overview
Send feedback
2
3
4
5
6
7
8
9
10
11
12:A
B
C
D
E
F
G
H
Pyridoxine
L-Arginine
Glutamic Acid
L-Tyrosine
L-Proline
L-Methionine
Niacin
Folic Acid
Serine
CCDC92
SERPINE1
Caffeine
Vitamin A
APOE
RHOA
ICAM1
SCARB1
AKT1
FN1
LIPA
MAT2A
SMAD3
coronary artery disease
COL4A1
LPL
ATP1B1
HTRA1
ARHGEF12
EDN1
IGF2R
JCAD
SMARCA4
IL1B
CCL2
KDELR2
NOS3
Vitamin E
VEGFA
© 2021 SPOKE
Terms
10,000 ft

The July/Aug 2022 special issue in *IEEE Computer Graphics and Applications* on "Multi-Level Graph Representations for Big Data in Science"

Articles due for review:
December 29, 2021

Guest editors:

- Katy Börner, Indiana University, Bloomington, US
- Stephen G. Kobourov, University of Arizona, Tucson, US

https://www.computer.org/digital-library/magazines/cg/call-for-papers-special-issue-on-multi-level-graph-representations-for-big-data-in-science

Call for Papers: Special Issue on Multi-level Graph Representations for Big Data in Science

CG&A seeks submissions for this upcoming special issue.

For centuries, cartographic maps have guided human exploration. While being rather imperfect initially, they helped explorers find promised lands and return home safely. Recent advances in data, algorithms, and computing infrastructures make it possible to map humankind's collective scholarly knowledge and technology expertise by using topic maps on which "continents" represent major areas of science (e.g., mathematics, physics, or medicine) and zooming reveals successively more detailed subareas. Basemaps of science and technology are generated by analyzing citations links between millions of publications and/or patents. "Data overlays" (e.g., showing all publications by one scholar, institution, or country or the career trajectory of a scholar as a pathway) are generated by science-locating relevant publication records based on topical similarity. Despite the demonstrated utility of such maps, current approaches do not scale to the hundreds of millions of data records now available. The main challenge is designing efficient and effective methods to visualize and interact with more than 100 million scholarly publications at multiple levels of resolution.

This special issue invites researchers in cartography, data visualization, science of science, graph drawing, and other domains to submit novel and promising new research on graph mining and layout algorithms and their application to the development of science mapping standards and services. Topics of interest include:

- Science of science user needs and applications
- Efficient multi-level graph algorithms
- Network visualizations
- Effective user interfaces to large-scale data visualizations

Deadlines

Submissions due: 29 December 2021
Preliminary notification: 2 March 2022
Revisions due: 6 April 2022
Final notification: 11 May 2022
Final version due: 25 May 2022
Publication: July/August 2022

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

HuBMAP: Mapping 30+ Trillion Cells

Michael P. Snyder, et al. 2019. The human body at cellular resolution: The NIH Human Biomolecular Atlas Program. *Nature*. 574, p. 187-192.
https://www.nature.com/articles/s41586-019-1629-x.pdf

HuBMAP

https://commonfund.nih.gov/HuBMAP

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

Goals

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

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

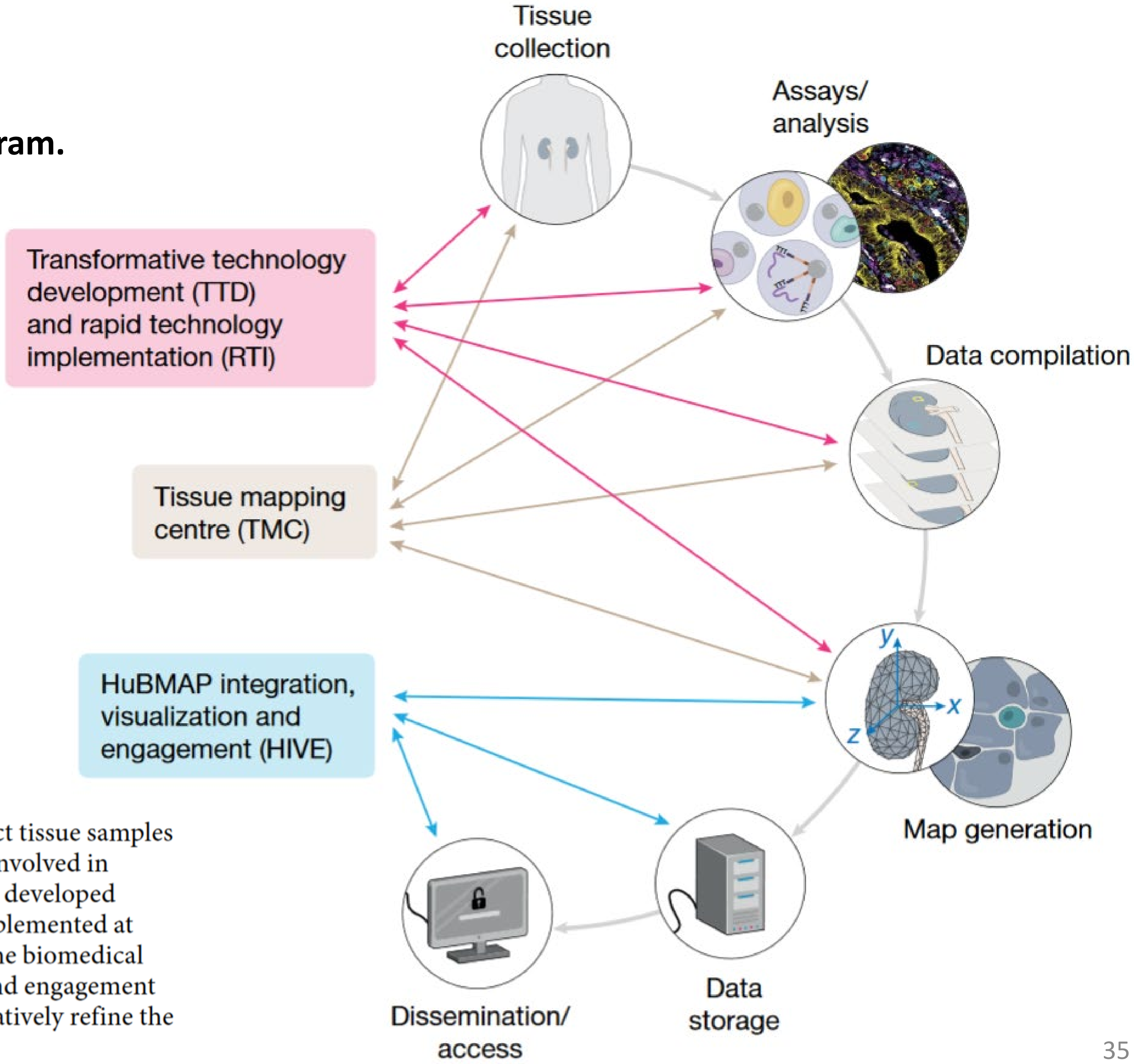


Fig. 1 | The HubMAP consortium. The TMCs will collect tissue samples and generate spatially resolved, single-cell data. Groups involved in TTD and RTI initiatives will develop emerging and more developed technologies, respectively; in later years, these will be implemented at scale. Data from all groups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

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

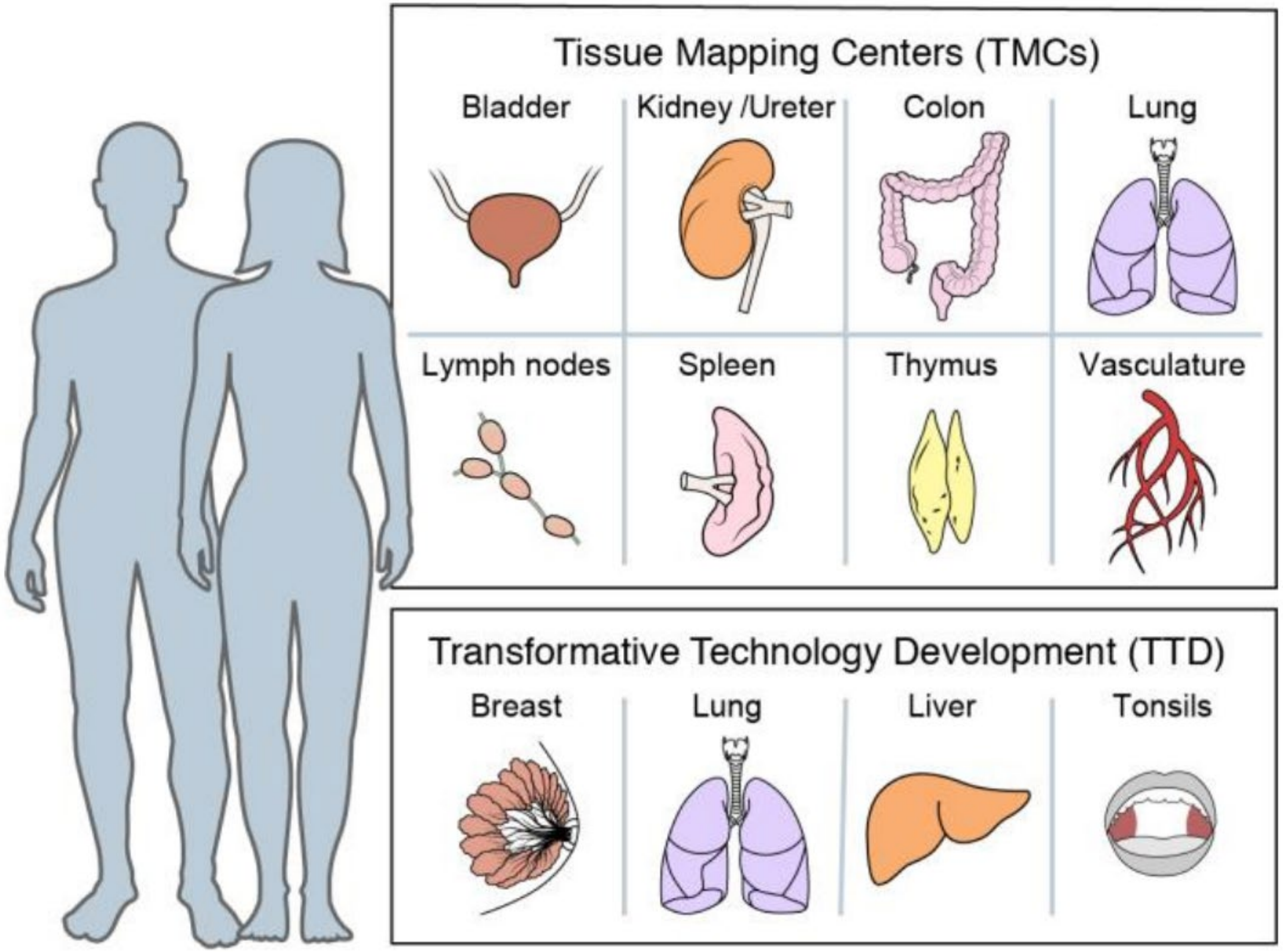


Fig. 2 | Key tissues and organs initially analysed by the consortium. Using innovative, production-grade ('shovel ready') technologies, HuBMAP TMCs will generate data for single-cell, three-dimensional maps of various human tissues. In parallel, TTD projects (and later RTI projects) will refine assays and analysis tools on a largely distinct set of human tissues. Samples from individuals of both sexes and different ages will be studied. The range of tissues will be expanded throughout the program.

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

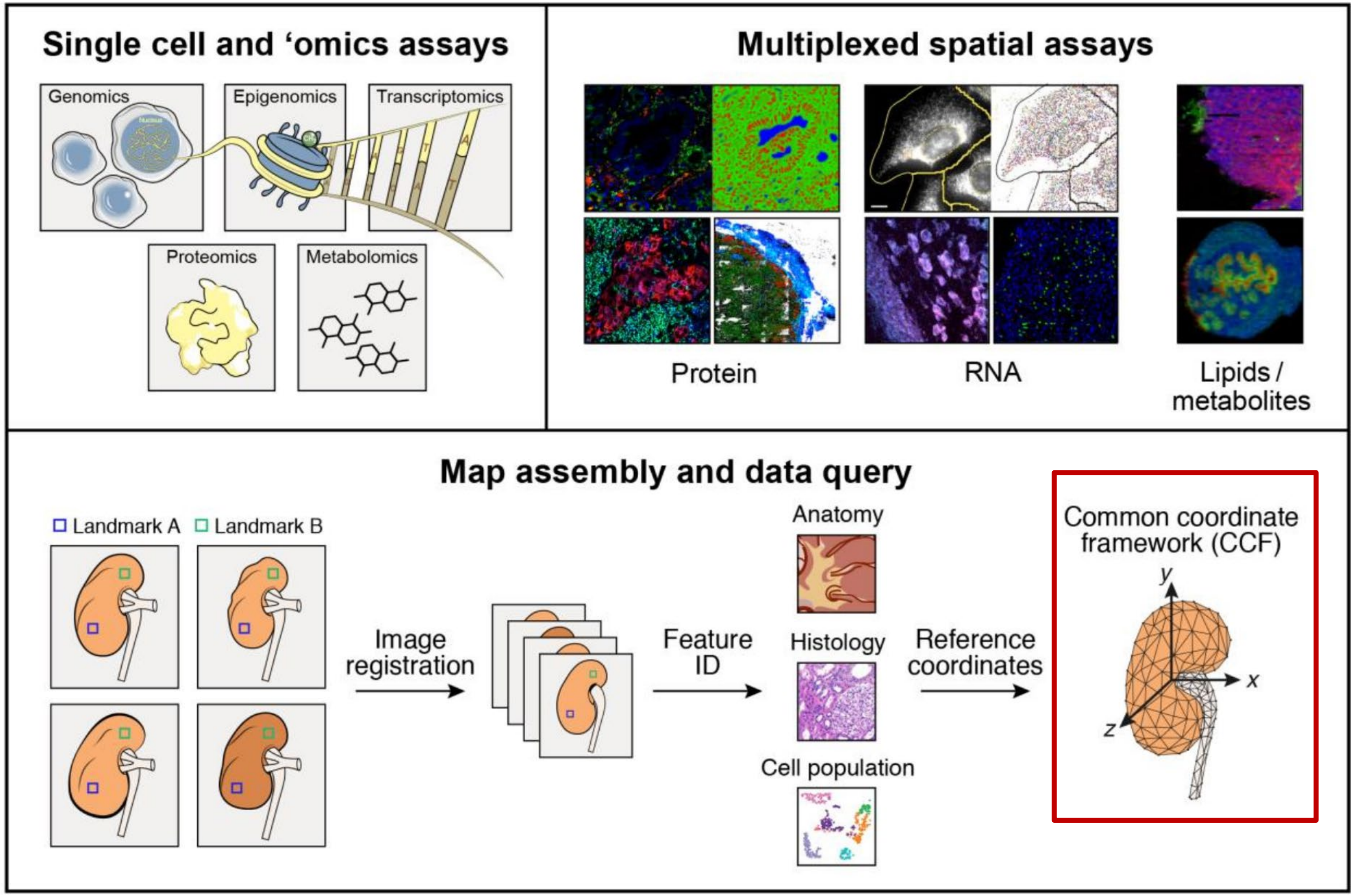
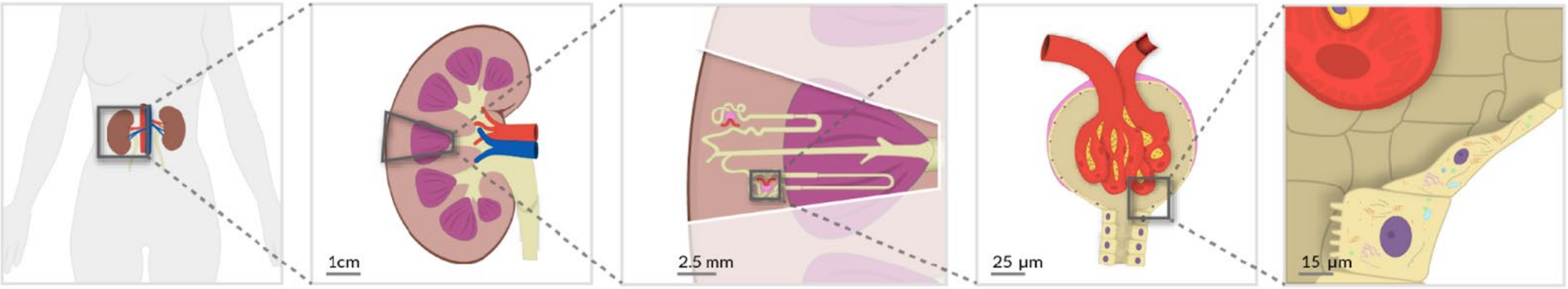


Fig. 3 | Map generation and assembly across cellular and spatial scales. HuBMAP aims to produce an atlas in which users can refer to a histological slide from a specific part of an organ and, in any given cell, understand its contents on multiple 'omic levels—genomic, epigenomic, transcriptomic, proteomic, and/or metabolomic. To achieve these ends, centres will apply a combination of imaging, 'omics and mass spectrometry techniques to specimens collected in a reproducible manner from specific sites in the body. These data will be then be integrated to arrive at a high-resolution, high-content three-dimensional map for any given tissue. To ensure inter-individual differences will not be confounded with collection heterogeneity, a robust CCF will be developed.

CCF Requirements

The CCF must capture major **anatomical structures, cell types, and biomarkers** and their interrelations across **multiple levels of resolution.**

It should be **semantically explicit** (using existing ontologies, e.g., Uberon, CL) and **spatially explicit** (e.g., using 3D reference organs for registration and exploration).



Body
- Body
- Kidney (Left, Right)
- Aorta
- Renal artery
- Renal vein
- Ureter

Organ
- Renal capsule
- Renal pyramid
- Renal cortex
- Renal medulla
- Renal calyx
- Renal pelvis

Functional Tissue Unit
- Nephron
- Renal corpuscle
- Proximal convoluted tubule
- Loop of Henle
- Distal convoluted tubule
- Connecting tubule
- Collecting duct

FTU Sub-structure(s)
- Bowman's capsule
- Glomerulus
- Efferent arteriole
- Afferent arteriole

Cellular
- Parietal epithelial cell
- Capillary endothelial cell
- Mesangial cell
- Podocyte

ASCT+B Tables

Anatomical Structures (AS), Cell Types (CT), and Biomarkers (B) or ASCT+B tables aim to capture the partonomy of anatomical structures, cell types, and major biomarkers (e.g., gene, protein, lipid or metabolic markers).

ASCT Table

Structure/Region	Sub structure/Sub region	Cell Type
Renal Corpuscle	Bowman's (glomular) Capsule/parietal layer	Parietal epithelial Cell
	Bowman's (glomular) Capsule/visceral layer	Podocyte
	Glomerular Tuft	Capillary Endothelial Cell
		Mesangial Cell
Tubules	Proximal Tubule	Proximal Tubule Epithelial Cell (general)
		Proximal Convoluted Tubule Epithelial Cell Segment 1
		Proximal Tubule Epithelial Cell Segment 2
		Proximal Tubule Epithelial Cell Segment 2
	Loop of Henle, Thin Limb	Descending Thin Limb Cell (general)
		Ascending Thin Limb Cell (general)
	Loop of Henle, Thick Limb	Thick Ascending Limb Cell (general)
		Cortex-TAL Cell
		Medulla-TAL Cell
		TAL-Macula Densa Cell
	Distal Convolution	Distal Convoluted Tubule Cell (general)
		DCT Type 1 Cell
		DCT Type 2 Cell
	Connecting Tubule	Connecting Tubule Cell (general)
		CNT-Principal Cell

Ontology

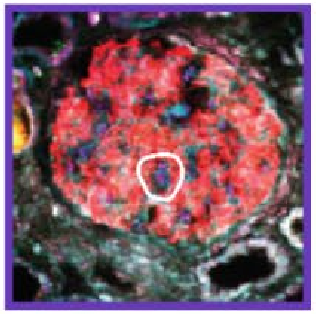
Anatomical Structures Partonomy

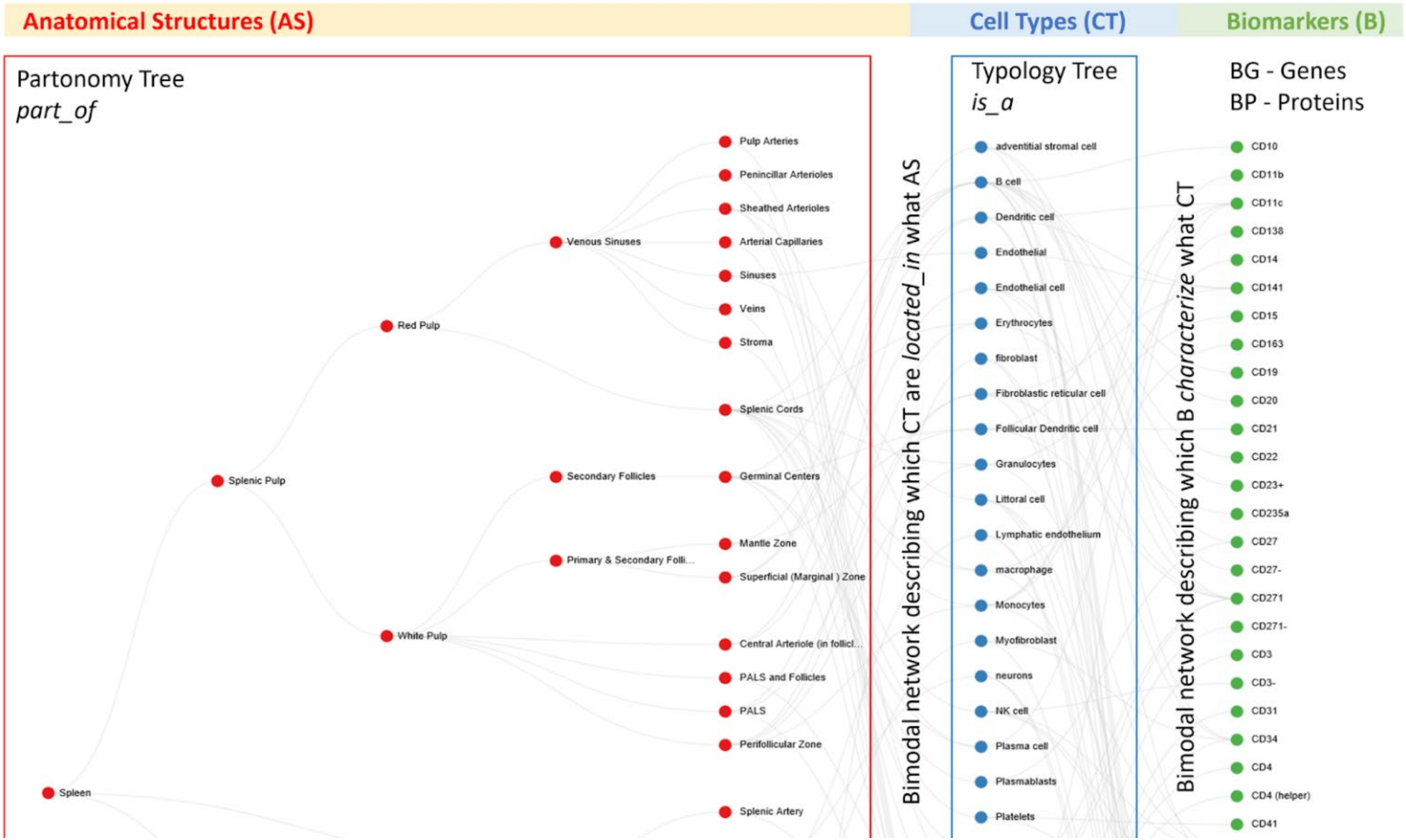
kidney
- kidney capsule
- cortex of kidney
  - outer cortex of kidney
- renal medulla

Cell Types Ontology

connective tissue cell
- pericyle cell
  - mesangial cell
    - extraglomerular mesangial cell
    - glomerular mesangial cell

3D Reference Object Library

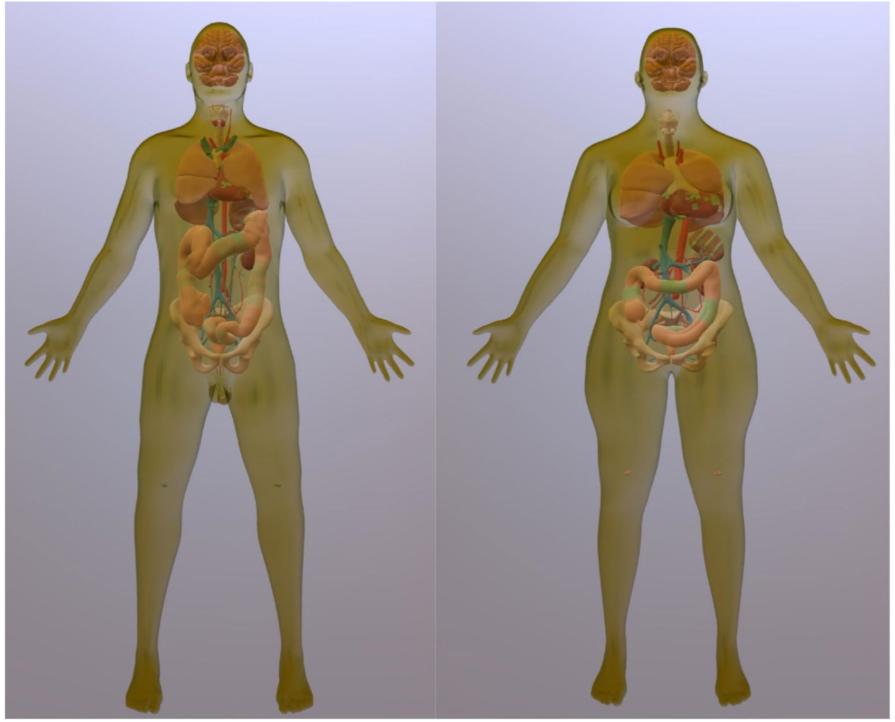



Anatomical Structures (AS)
Cell Types (CT)
Biomarkers (B)
Partonomy Tree
part_of
Spleen
Splenic Pulp
Red Pulp
White Pulp
Venous Sinuses
Secondary Follicles
Primary & Secondary Folli...
Pulp Arteries
Penincillar Arterioles
Sheathed Arterioles
Arterial Capillaries
Sinuses
Veins
Stroma
Splenic Cords
Germinal Centers
Mantle Zone
Superficial (Marginal) Zone
Central Arteriole (in follicl...
PALS and Follicles
PALS
Perifollicular Zone
Splenic Artery
Bimodal network describing which CT are located_in what AS
Typology Tree
is_a
adventitial stromal cell
B cell
Dendritic cell
Endothelial
Endothelial cell
Erythrocytes
fibroblast
Fibroblastic reticular cell
Follicular Dendritic cell
Granulocytes
Littoral cell
Lymphatic endothelium
macrophage
Monocytes
Myofibroblast
neurons
NK cell
Plasma cell
Plasmablasts
Platelets
Bimodal network describing which B characterize what CT
BG - Genes
BP - Proteins
CD10
CD11b
CD11c
CD138
CD14
CD141
CD15
CD163
CD19
CD20
CD21
CD22
CD23+
CD235a
CD27
CD27-
CD271
CD271-
CD3
CD3-
CD31
CD34
CD4
CD4 (helper)
CD41

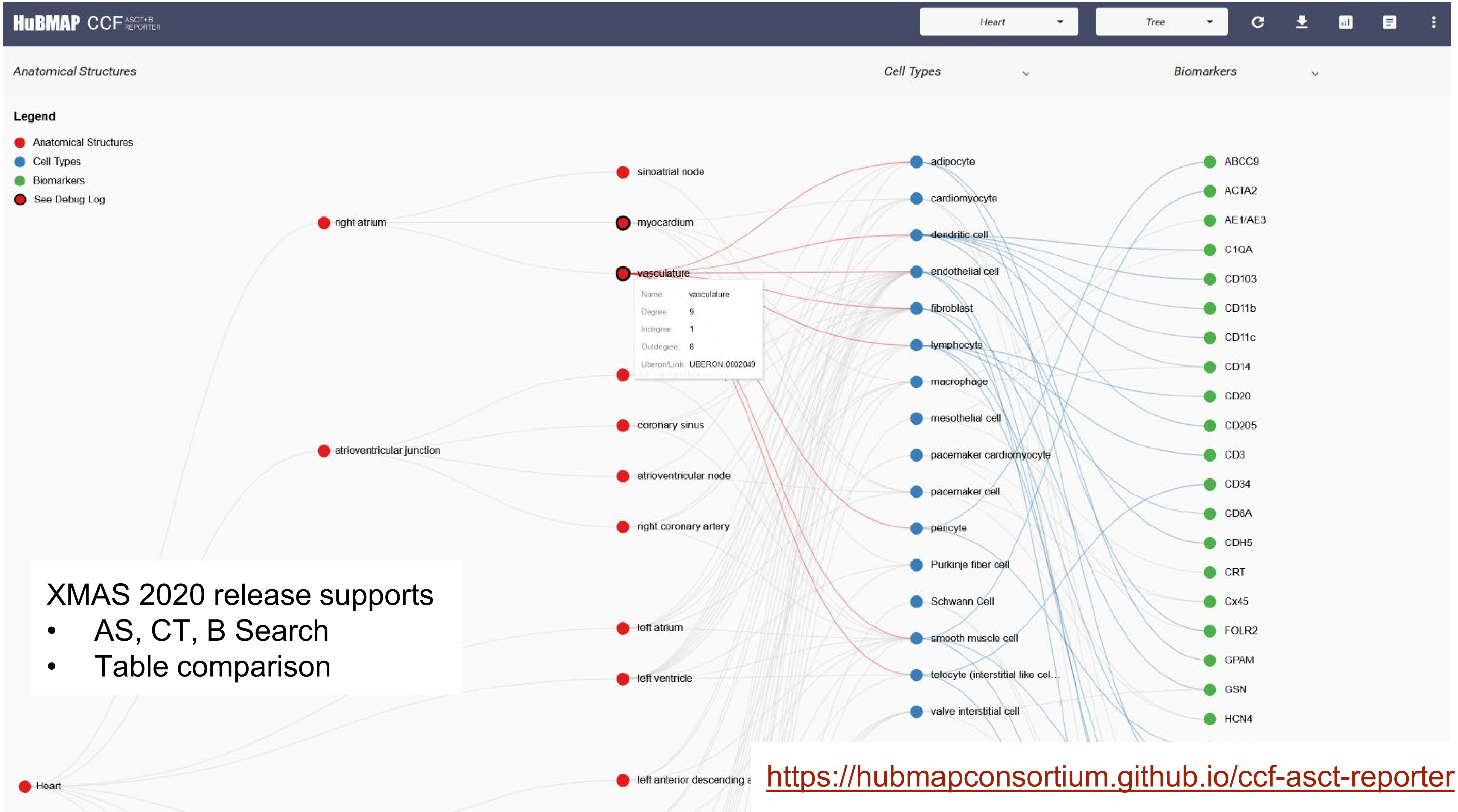
	3D Ref. Organ				ASCT+B Table					
Organ	#AS M-L	#AS F-L	#AS M-R	#AS F-R	#AS	#CT	#B	#AS-AS (part_of)	#CT-AS (located_in)	#B-CT (characterizes)
BM & Blood / Pelvis	23	23			14	46	202	24	97	296
Brain	141	141			187	127	29	187	127	36
Heart	39	46			50	25	48	57	164	78
Intestine, Large	10	10			66	69	89	409	1410	192
Kidney	38	44	39	41	64	64	129	63	58	215
Lung	74	74			91	85	174	108	123	296
Lymph Nodes	7	7	7	7	40	49	161	60	117	342
Skin	1	1			16	42	70	17	19	105
Spleen	8	8			46	66	0	68	172	0
Thymus	2	2			18	46	55	20	103	64
Vasculature	84	85			869	2	1	868	606	2
Totals	**427**	**441**	**46**	**48**	**1461**	**621**	**958**	**1881**	**2996**	**1626**

https://hubmapconsortium.github.io/ccf/pages/ccf-anatomical-structures.html

Male *Female*



https://hubmapconsortium.github.io/ccf/pages/ccf-3d-reference-library.html (NLM VH organs)
https://community.brain-map.org/t/allen-human-reference-atlas-3d-2020-new/ (brain)
https://www3.cs.stonybrook.edu/~ari/ (male colon)



XMAS 2020 release supports

- AS, CT, B Search
- Table comparison

https://hubmapconsortium.github.io/ccf-asct-reporter

Document the tissue extraction site by registering tissue blocks within a 3D reference organ.

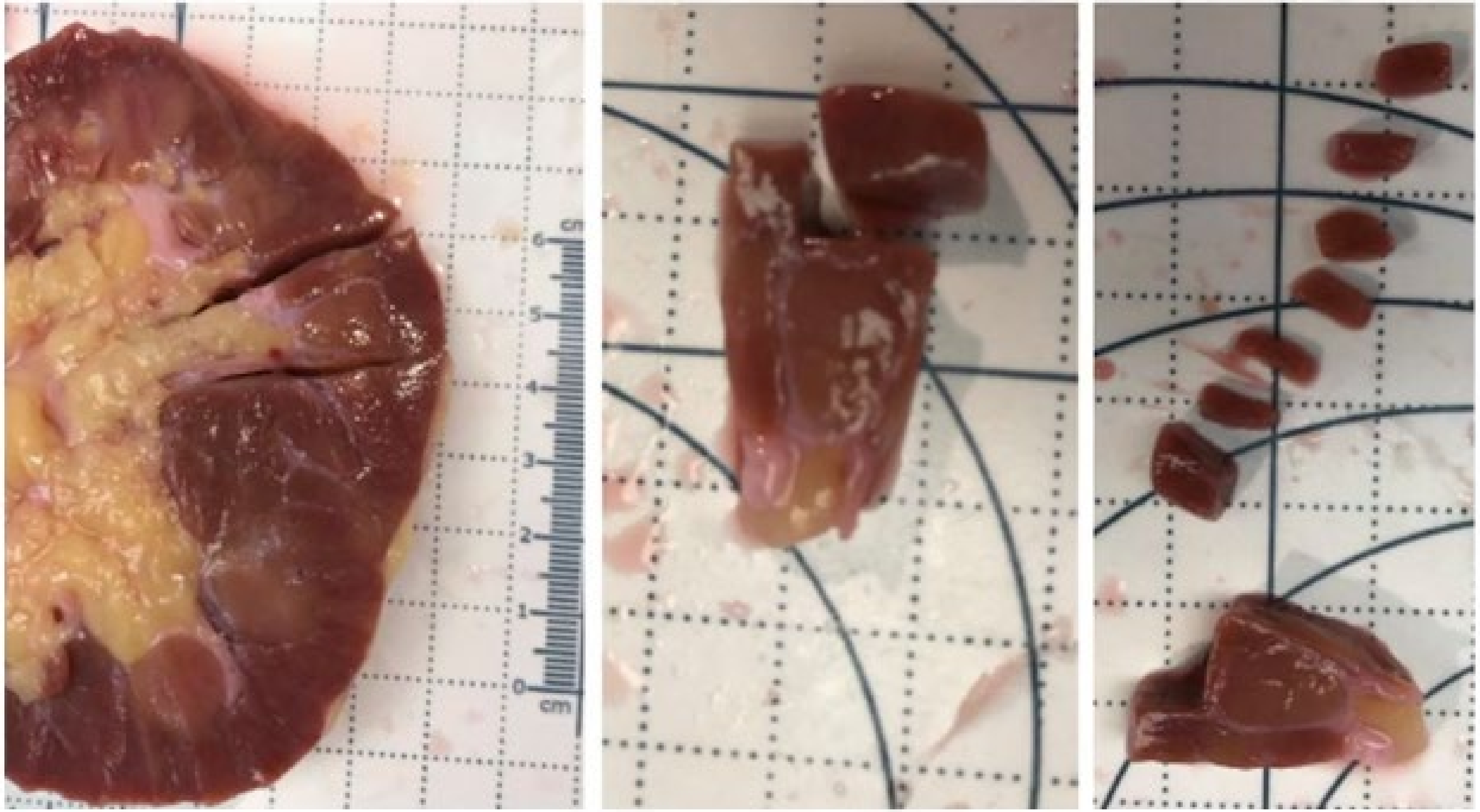
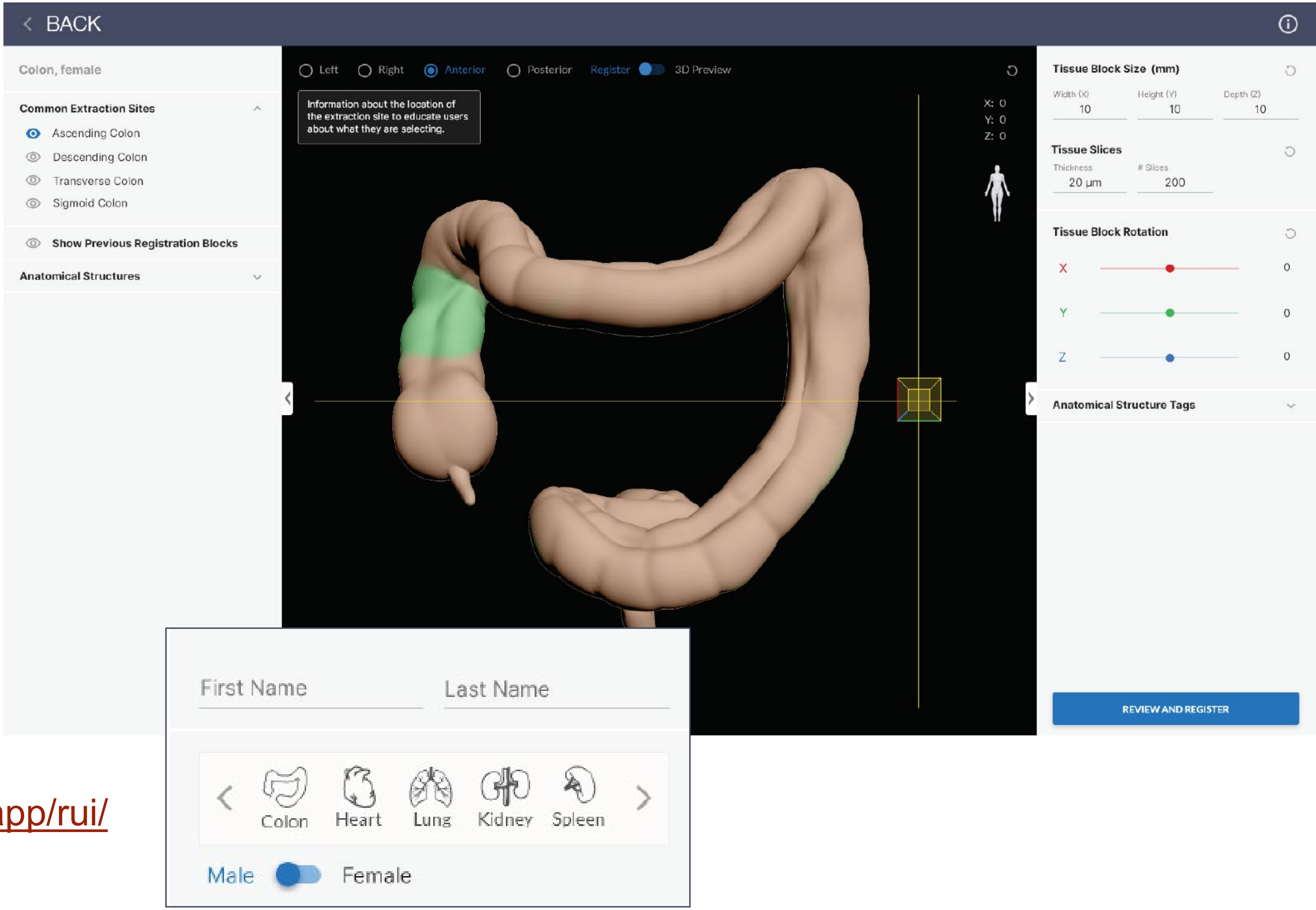


Image provided by Sanjay Jain, TMC-UCSD

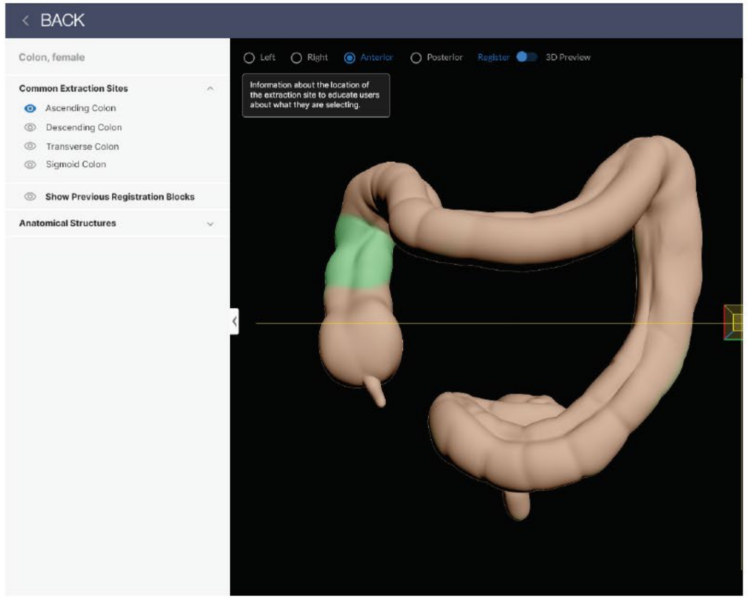
CCF Registration User Interface (RUI) v1.0.0

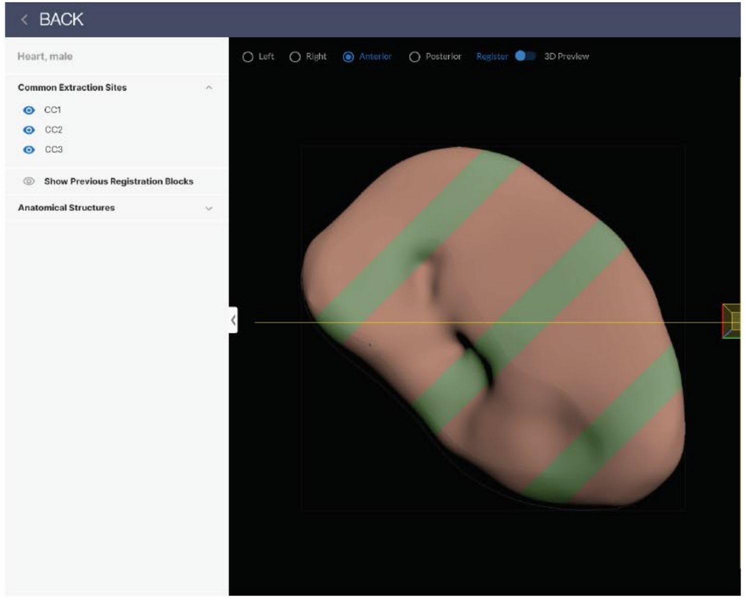
New Features:

- Organ carousel with 4 reference organs
- Support for tissue extraction sites
- Expanded ontology
- Semantic annotation via collision detection & manual annotation
- Support for non-HuBMAP usage



https://hubmap-ccf-ui.netlify.app/rui/





Kidney

- Bisection Line

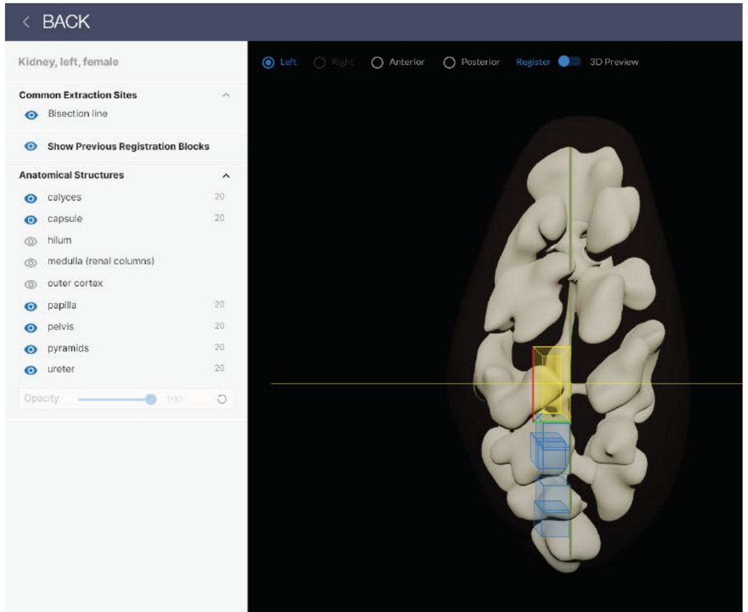
Spleen

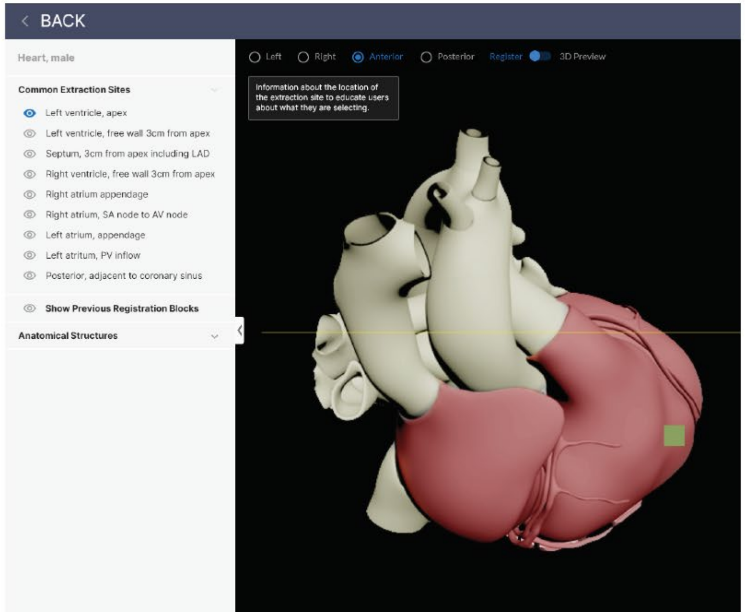
- CC1
- CC2
- CC3

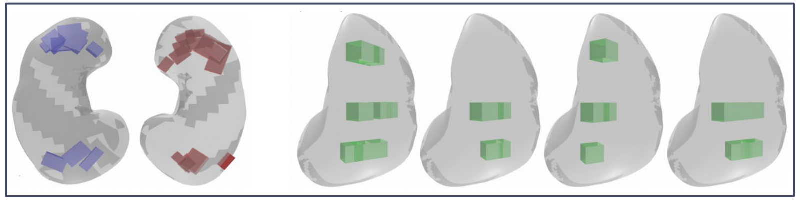
Colon

- Ascending Colon
- Descending Colon
- Transverse Colon
- Sigmoid Colon

Heart	Extraction Site Mapping
Left atrium, appendage	7
Left atrium, PV inflow	8
Left ventricle, apex	1
Left ventricle, free wall 3cm from apex	2
Septum, 3cm from apex including LAD	3
Posterior, adjacent to coronary sinus	9
Right atrium appendage	5
Right atrium, AV (atrioventricular) node	6a
Right atrium, SA (sinoatrial) node	6b
Right ventricle, free wall 3cm from apex	4

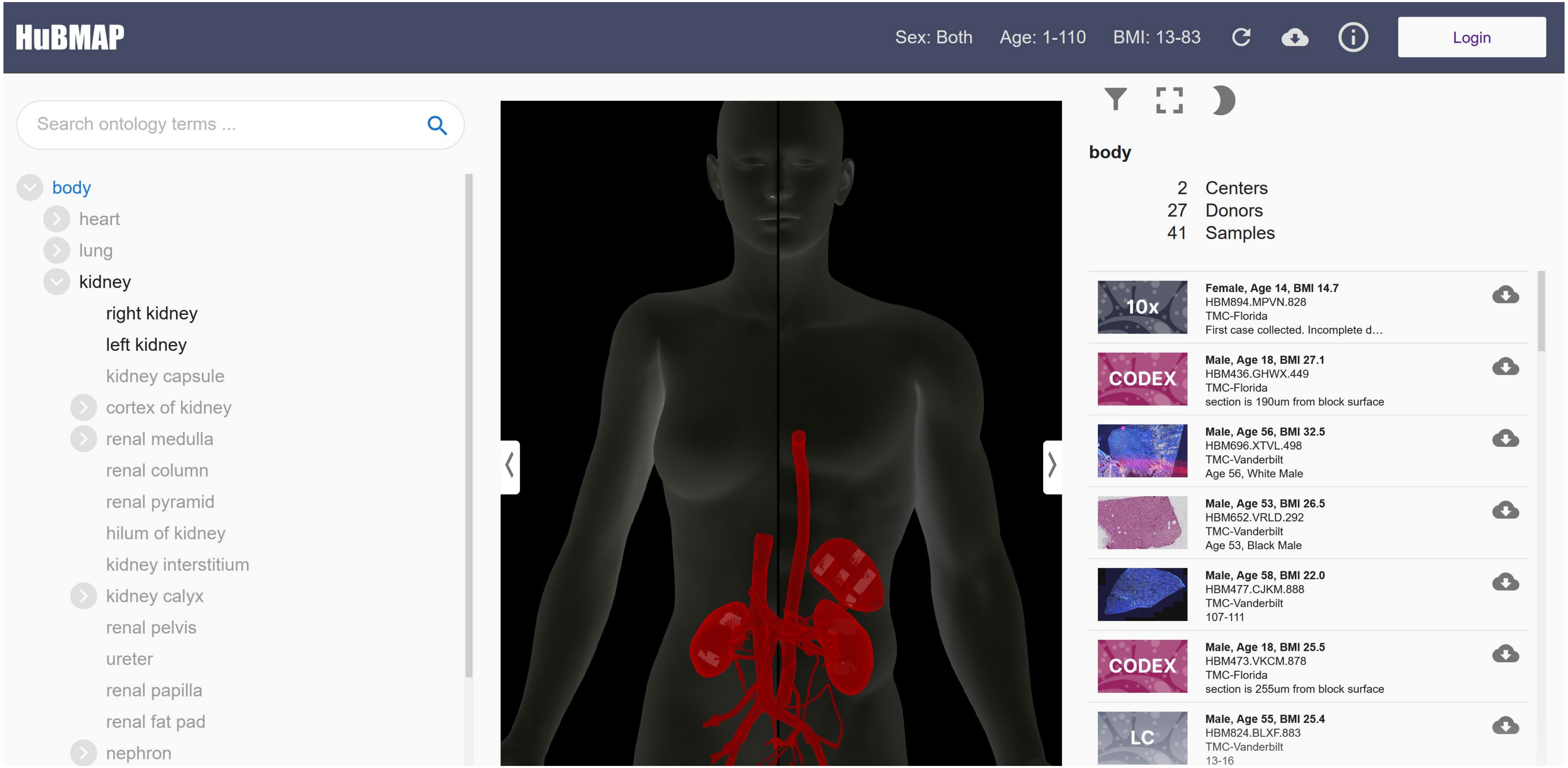




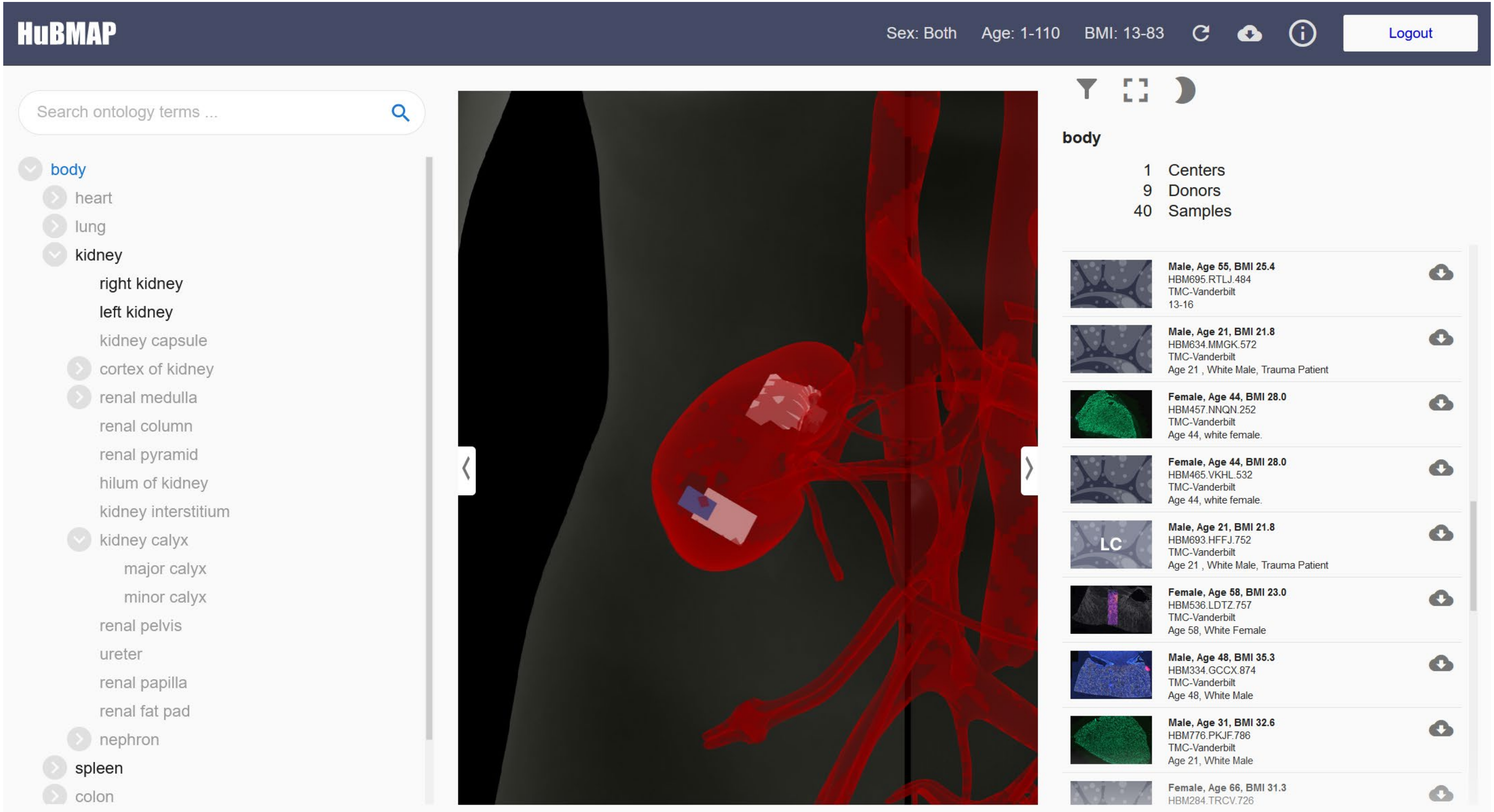


For the first HuBMAP portal release, 48 tissue blocks were registered.

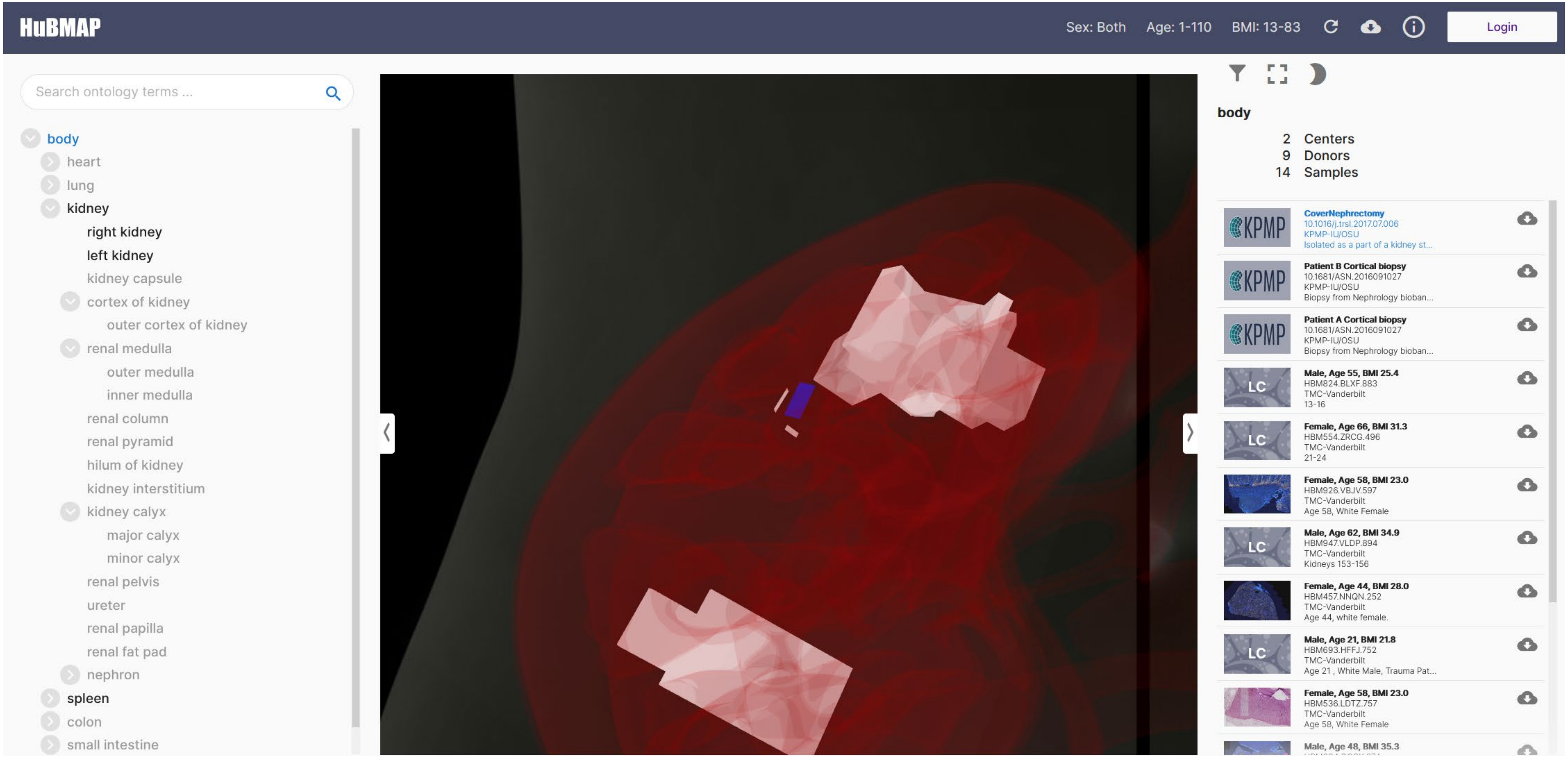
CCF Exploration User Interface (EUI)



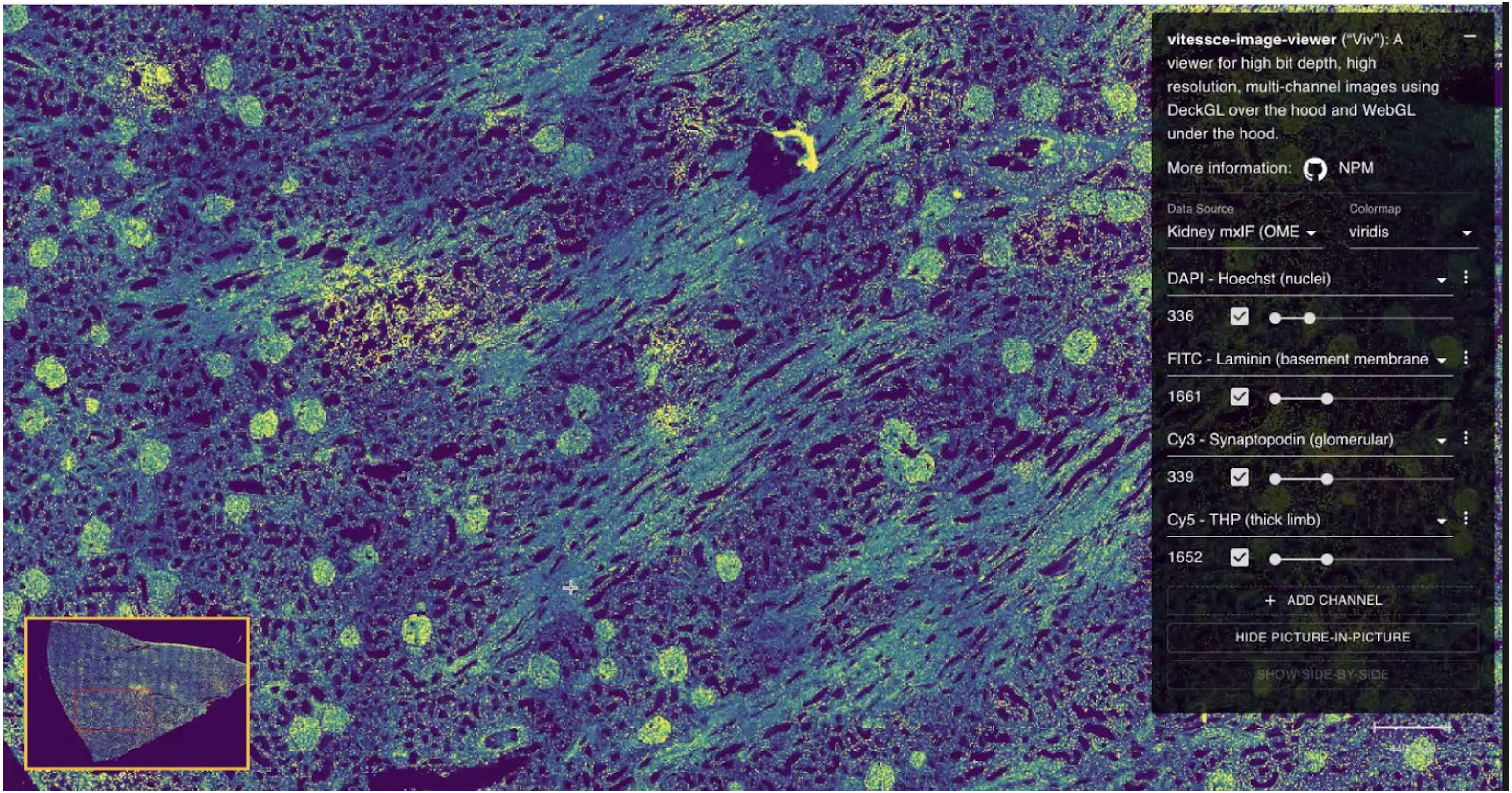
https://portal.hubmapconsortium.org/ccf-eui



HuBMAP
Sex: Both
Age: 1-110
BMI: 13-83
Logout
Search ontology terms ...
body
heart
lung
kidney
right kidney
left kidney
kidney capsule
cortex of kidney
renal medulla
renal column
renal pyramid
hilum of kidney
kidney interstitium
kidney calyx
major calyx
minor calyx
renal pelvis
ureter
renal papilla
renal fat pad
nephron
spleen
colon
body
1 Centers
9 Donors
40 Samples
Male, Age 55, BMI 25.4
HBM695.RTLJ.484
TMC-Vanderbilt
13-16
Male, Age 21, BMI 21.8
HBM634.MMGK.572
TMC-Vanderbilt
Age 21 , White Male, Trauma Patient
Female, Age 44, BMI 28.0
HBM457.NNQN.252
TMC-Vanderbilt
Age 44, white female.
Female, Age 44, BMI 28.0
HBM465.VKHL.532
TMC-Vanderbilt
Age 44, white female.
LC
Male, Age 21, BMI 21.8
HBM693.HFFJ.752
TMC-Vanderbilt
Age 21 , White Male, Trauma Patient
Female, Age 58, BMI 23.0
HBM536.LDTZ.757
TMC-Vanderbilt
Age 58, White Female
Male, Age 48, BMI 35.3
HBM334.GCCX.874
TMC-Vanderbilt
Age 48, White Male
Male, Age 31, BMI 32.6
HBM776.PKJF.786
TMC-Vanderbilt
Age 21, White Male
Female, Age 66, BMI 31.3
HBM284.TRCV.726



Register your data via https://hubmap-ccf-ui.netlify.app/rui/ so it can be spatially/semantically explored in EUI.



http://gehlenborglab.org/research/projects/vitessce/

VH Massive Open Online Course (VHMOOC)

Goals

- Communicate tissue data acquisition and analysis,
- Demonstrate single-cell analysis and CCF mapping techniques, and
- Introduce major features of the HuBMAP portal.

Learning modules come with

- Videos (incl. interviews, tool demos)
- Hands-on exercises
- Self-quizzes

HuBMAP Visible Human MOOC (VHMOOC)

Started Aug 4, 2020

GO TO CANVAS COURSE

You are enrolled.

INDIANA UNIVERSITY

Course Introduction

This 10h course introduces the HuBMAP project which aims to create an open, global reference atlas of the human body at the cellular level. Among others, the course describes the compilation and coverage of HuBMAP data, demonstrates new single-cell analysis and mapping techniques, and introduces major features of the HuBMAP portal.

Delivered entirely online, all coursework can be completed asynchronously to fit busy schedules. If you have questions or experience issues during registration, please email cnscntr@indiana.edu.

Learning Outcomes

- Theoretical and practical understanding of different single-cell tissue analysis techniques.
- Expertise in single-cell data harmonization used to federate data from different individuals analyzed using different technologies in diverse labs.
- Hands-on skills in the design and usage of semantic ontologies that describe human anatomy, cell types, and biomarkers (e.g., marker genes or proteins).
- Knowledge on the design and usage of a semantically annotated three-dimensional reference system for the healthy human body.
- An understanding of how the HuBMAP reference atlas might be used to understand human health but also to diagnose and treat disease.

Module Topics Include

- HuBMAP Overview: Project Goals, Setup, and Ambitions
- Tissue Data Acquisition and Analysis
- Biomolecular Data Harmonization
- Ontology, 3D Reference Objects, and User Interfaces
- HuBMAP Portal Design and Usage

Meet the Instructors

Katy Börner, Victor H. Yngve Distinguished Professor of Engineering and Information Science. Founding Director of the Cyberinfrastructure for Network Science Center at Indiana University.

Ellen M. Quardokus, staff in the Chemistry Department and research scientist, Cyberinfrastructure for Network Science Center, SICE with expertise in molecular biology, microscopy, anatomy, and interdisciplinary communication.

Andreas Bueckle, PhD Candidate in Information Science, performing research on information visualization, specifically virtual and augmented reality.

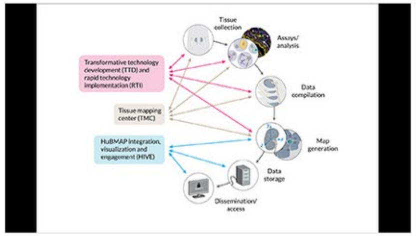
Length: 10 hours

Department: Cyberinfrastructure Network Science

Credit: None

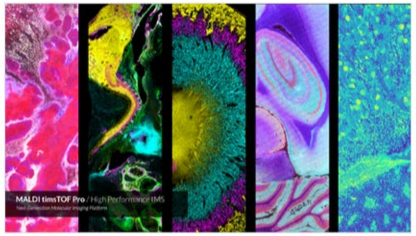
Audience: Biomedical students and professionals interested in single-cell tissue analysis and visualization

https://expand.iu.edu/browse/sice/cns/courses/hubmap-visible-human-mooc



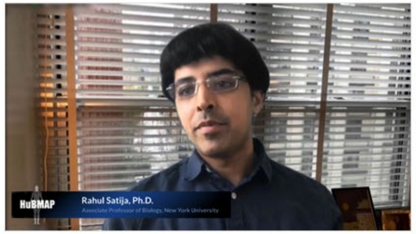
[HuBMAP Overview]

- Project Goals, Setup, and Ambitions



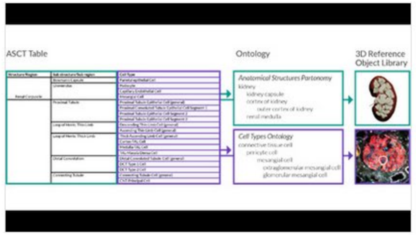
[Tissue Data Acquisition and Analysis]

- Behind the Scenes at Vanderbilt University



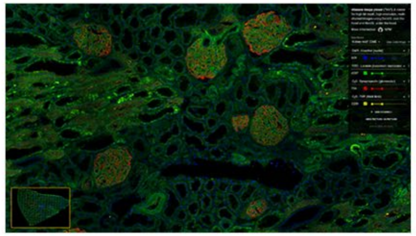
[Biomolecular Data Harmonization]

- An Introduction to Seurat



[CCF Ontology, 3D Reference Objects, and User Interfaces]

- Creating an Atlas of the Human Body

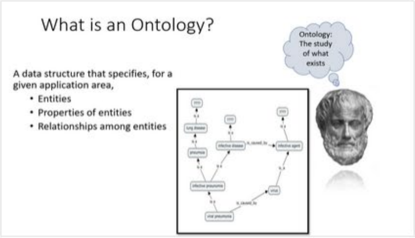


[Portal Design and Usage]

- Datasets and Software in the 1st HuBMAP Portal Release

[Open Consent Your Data]

- In Support of Research



[Ontologies 101]

- A gentle introduction on how to use ontologie the world.

[Anatomical Structures, Cell Types, and Biomarkers (ASCT+B) Tables]

- What are ASCT+B tables and how they are used.

Acknowledgements

HuBMAP Consortium (https://hubmapconsortium.org)

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

National Institutes of Health

Opportunity Pool Awardee
KIDNEY PRECISION MEDICINE PROJECT
kpmp.org

NSF

TMCs

Jeffrey Spraggins
TMC-Vanderbilt
Vanderbilt University

Sanjay Jain
TMC-UCSD
Washington University, St. Louis

Clive Wasserfall
TMC-UFL
University of Florida

Marda Jorgensen
TMC-UFL
University of Florida

3D Models

Kristen Browne
Medical Imaging and 3D Modeling Specialist
NIAID

MC-IU HIVE Team

Katy Börner
MC-IU PI
CNS Director

Griffin Weber
Assoc. Professor of Medicine
Harvard Medical School

Lisel Record
MC-IU PM
CNS Associate Director

Bruce Herr II
Sr. Systems Architect/PM

Ellen Quardokus
Sr. Research Analyst

Yingnan Ju
PhD Candidate

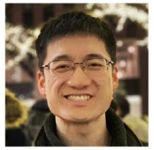
Andreas Bueckle
PhD Candidate

Leonard Cross
Sr. UX/UI Designer

Matthew Martindale
Center Assistant

Daniel Bolin
Software Developer

Adam Phillips
Software Developer



Edward Lu
Software Developer

Paul Hrishikesh
Research Assistant

Leah Scherschel
Research Assistant

Avinash Boppana
Research Consultant

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Data Visualization Literacy Framework

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

Data Visualization Literacy (DVL)

Data visualization literacy (ability to read, make, and explain data visualizations) requires:

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

Being able to "read and write" data visualizations is becoming as important as being able to read and write text. Understanding, measuring, and improving data and visualization literacy is important to strategically approach local and global issues.

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.

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

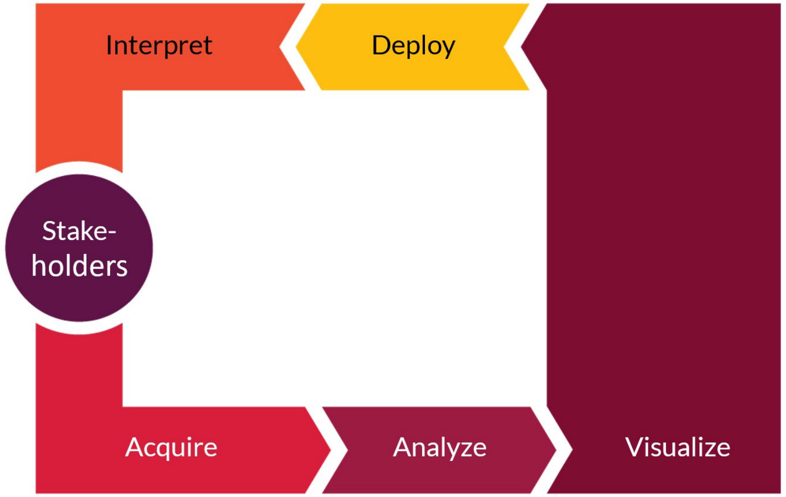
5. **Graphic Symbols**
   - geometric symbols
     point
     line
     area
     surface
     volume
   - linguistic symbols
     text
     numerals
     punctuation marks
   - pictorial symbols
     images
     icons
     statistical glyphs

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

DVL Workflow Process

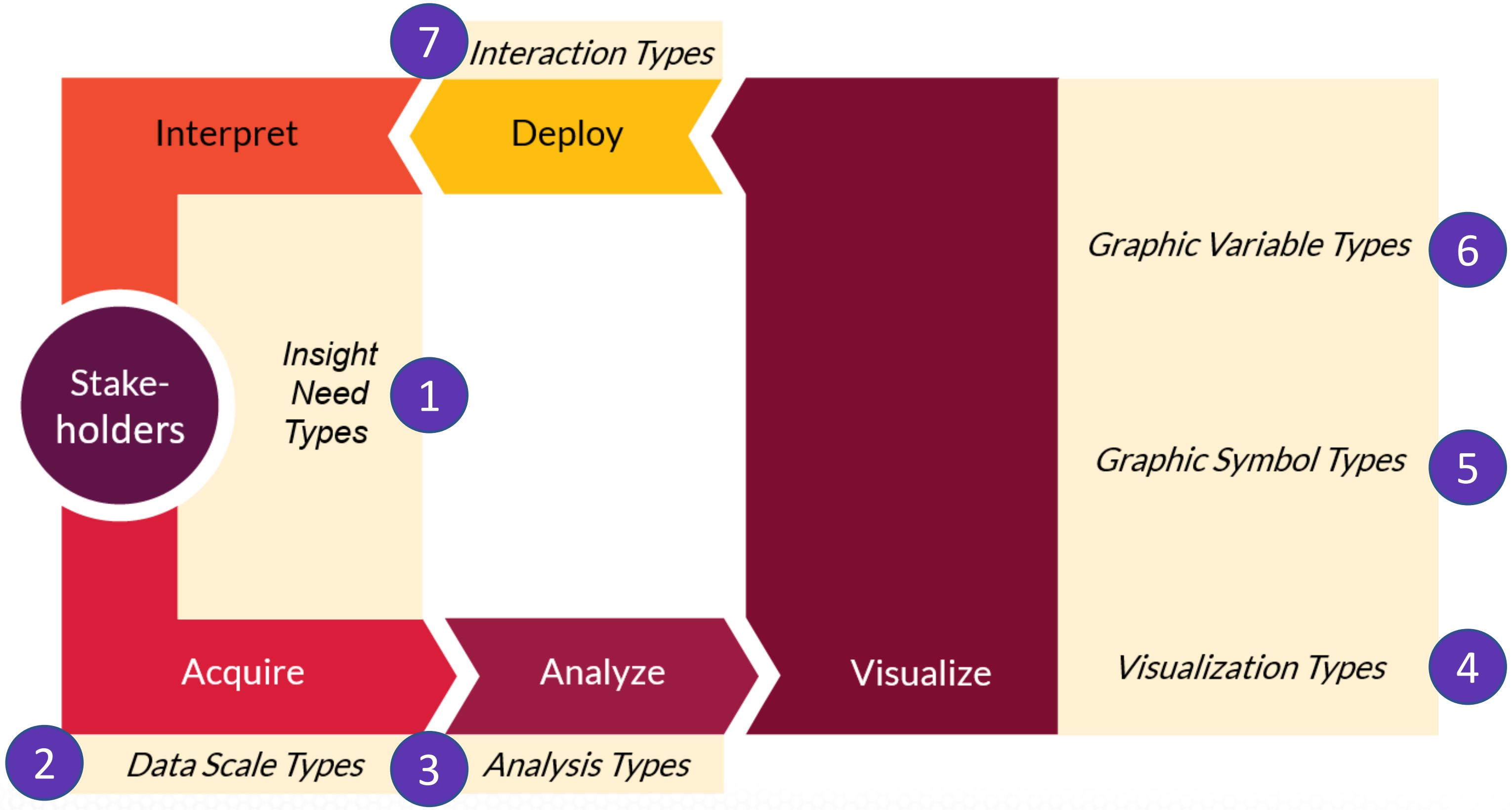
Defines 5 steps required to render data into insights.

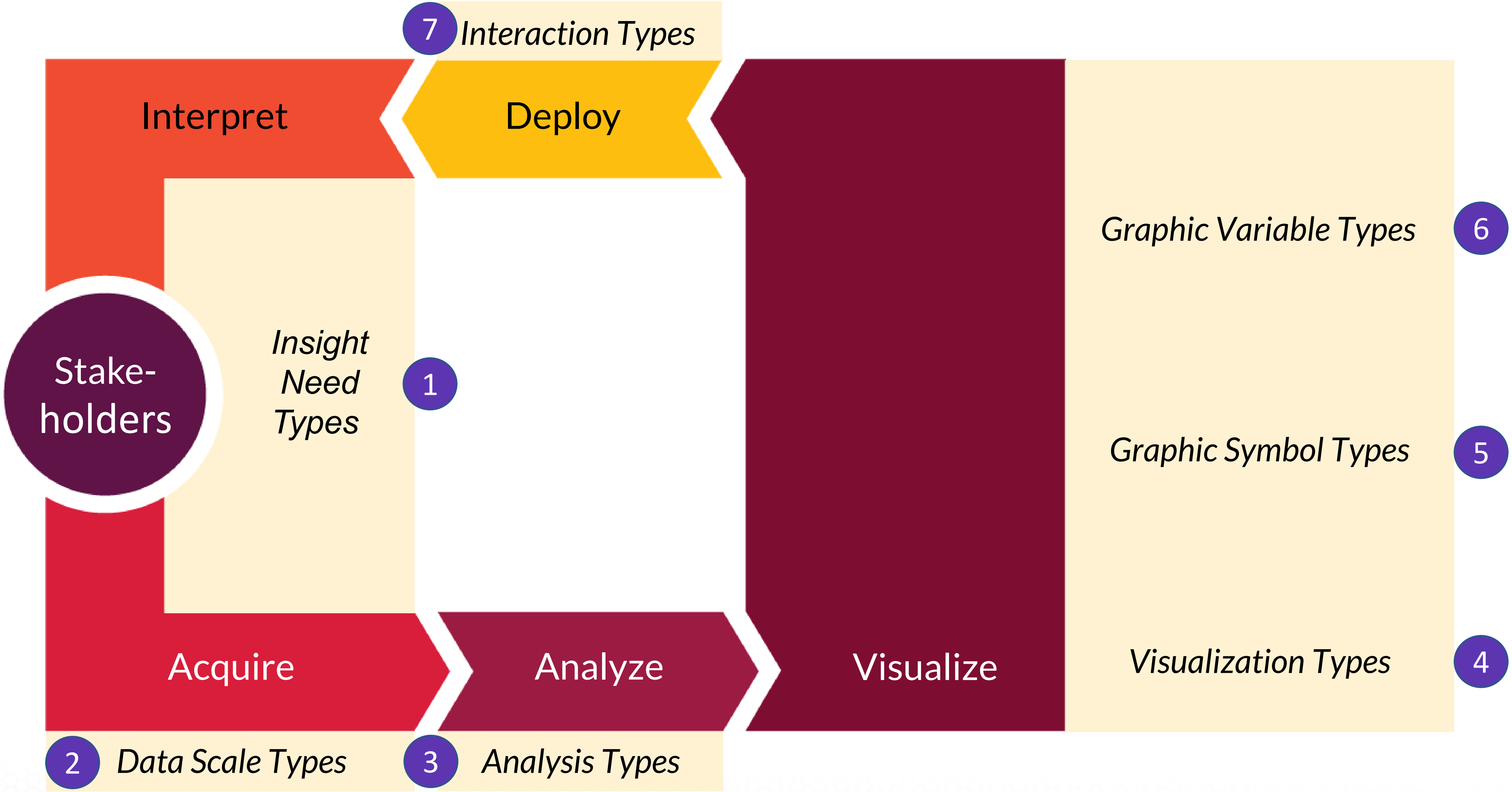


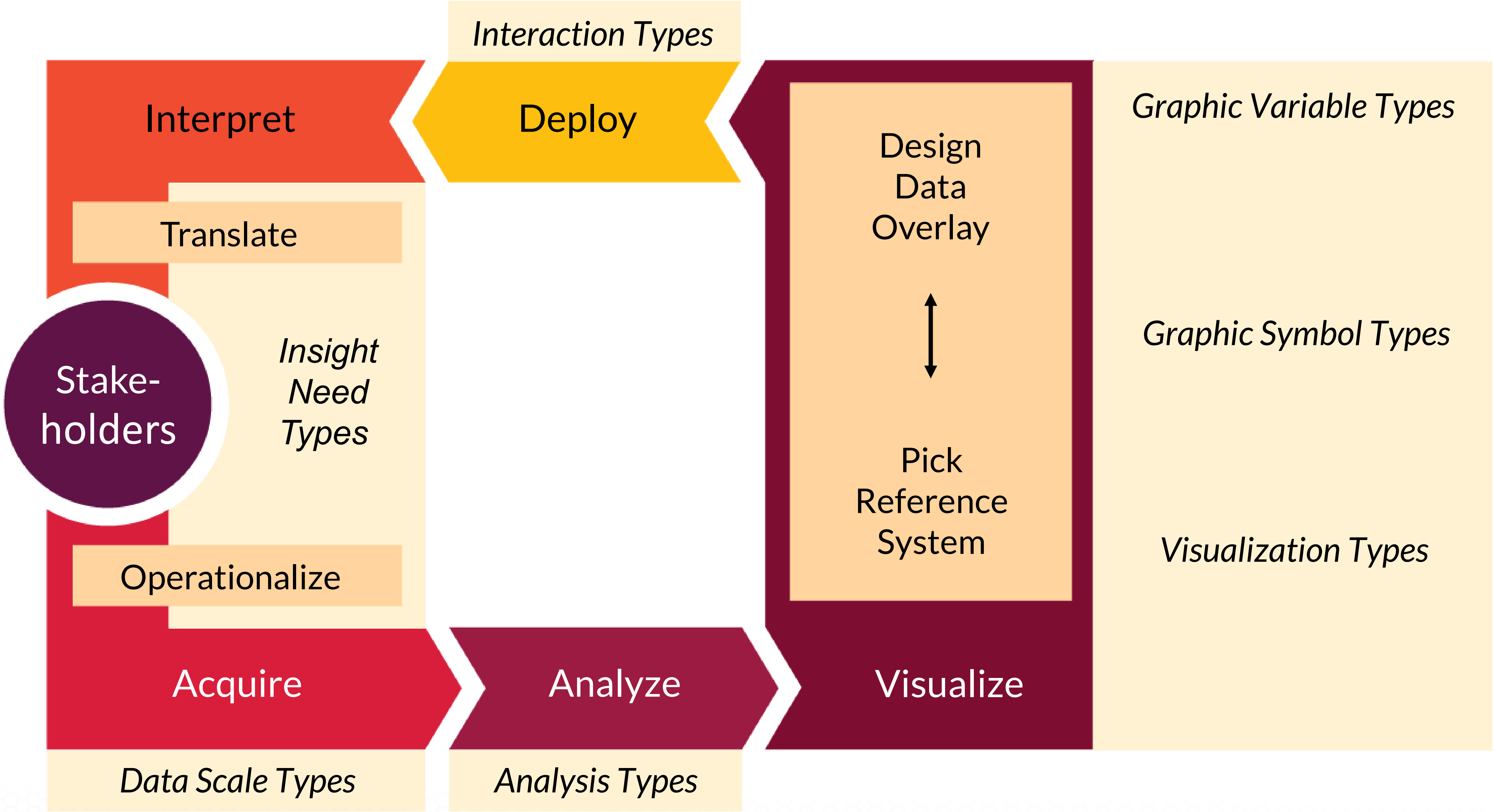
Data Visualization Literacy Framework (DVL-FW)

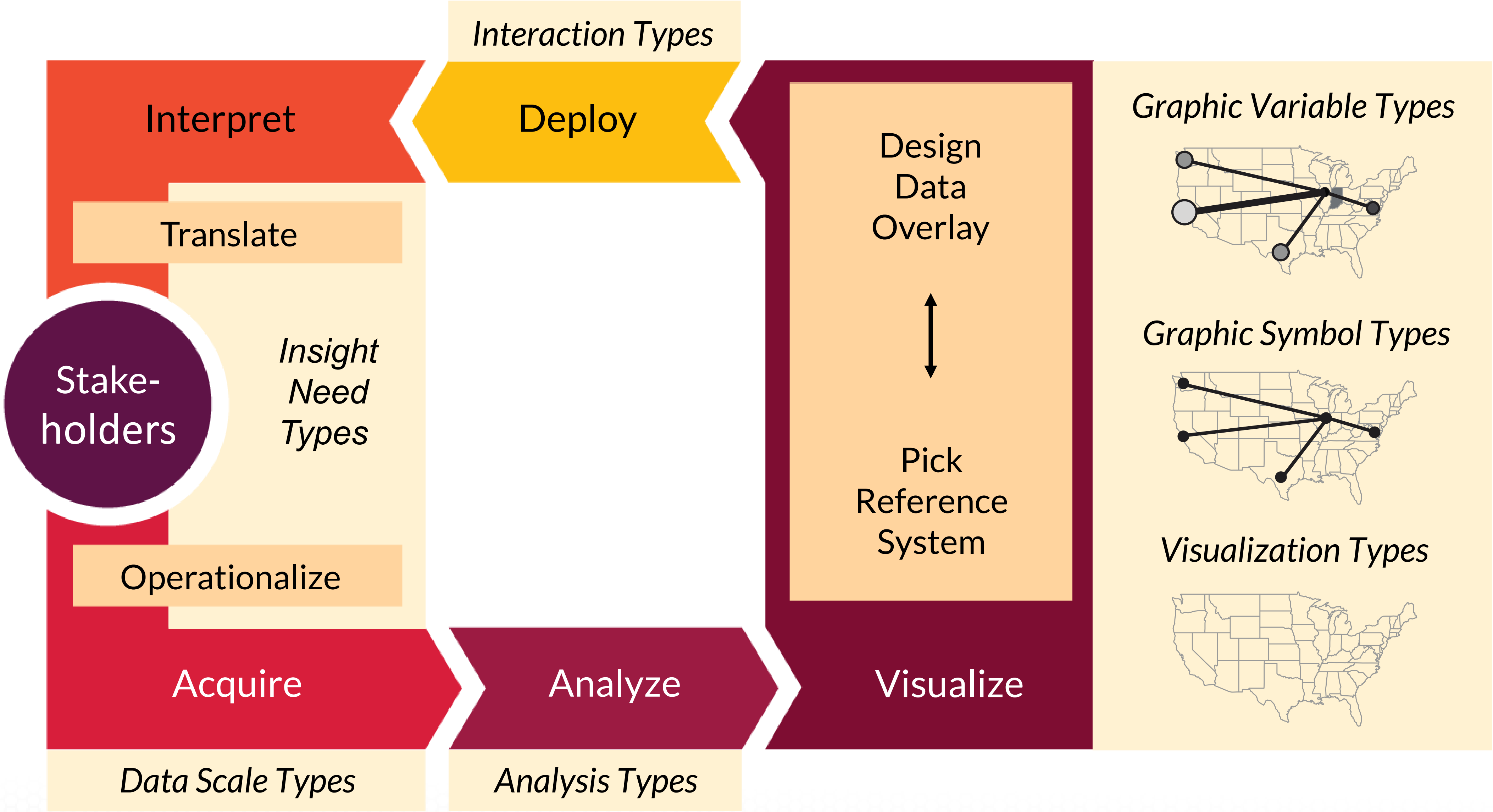
Consists of two parts that are interlinked:

DVL Typology +
DVL Workflow Process




7 Interaction Types
Interpret
Deploy
Stake-
holders
Insight
Need
Types
1
Graphic Variable Types
6
Graphic Symbol Types
5
Acquire
Analyze
Visualize
Visualization Types
4
2 Data Scale Types
3 Analysis Types

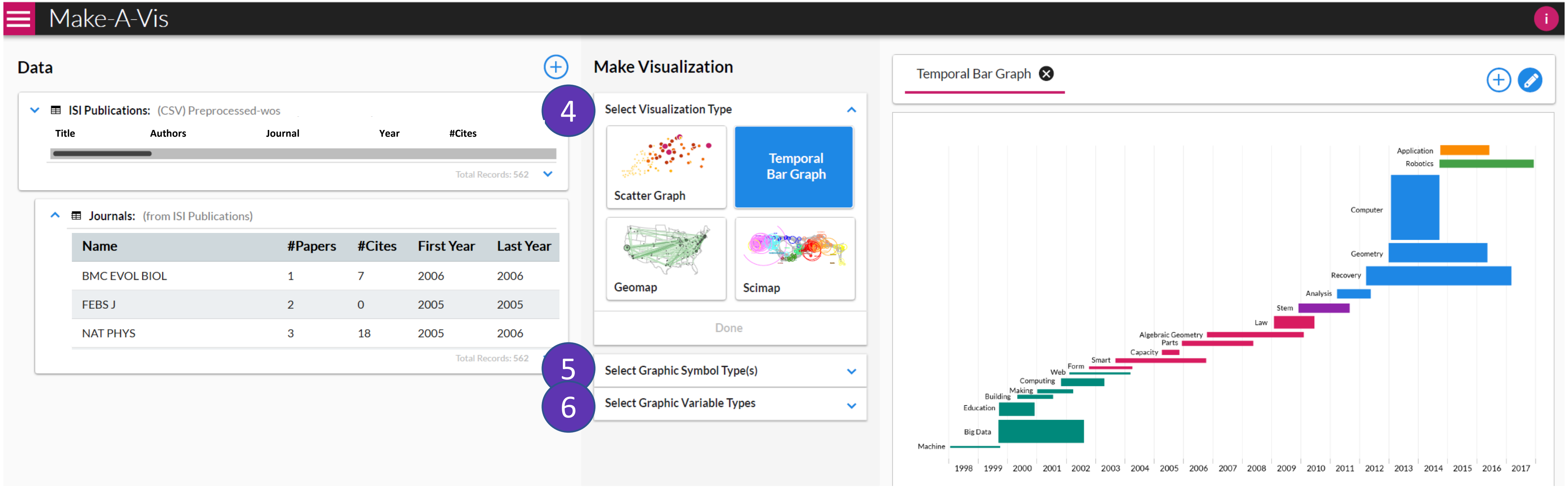

Interaction Types
Interpret
Deploy
Translate
Stake-holders
Insight Need Types
Operationalize
Design
Data
Overlay
Pick
Reference
System
Graphic Variable Types
Graphic Symbol Types
Visualization Types
Acquire
Analyze
Visualize
Data Scale Types
Analysis Types



Interaction Types
Interpret
Deploy
Translate
Stake-holders
Insight Need Types
Operationalize
Design Data Overlay
Pick Reference System
Graphic Variable Types
Graphic Symbol Types
Visualization Types
Acquire
Analyze
Visualize
Data Scale Types
Analysis 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

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

Typology of the Data Visualization Literacy Framework

1

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

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

Bertin, 1967	Wehrend & Lewis, 1996	Few, 2004	Yau, 2011	Rendgen & Wiedemann, 2012	Frankel, 2012	Tool: Many Eyes	Tool: Chart Chooser	Börner, 2014
selection	categorize			category				categorize/ cluster
order	rank	ranking					table	order/rank/ sort
	distribution	distribution					distribution	distributions (also outliers, gaps)
	compare	nominal comparison & deviation	differences		compare and contrast	compare data values	comparison	comparisons
		time series	patterns over time	time	process and time	track rises and falls over time	trend	trends (process and time)
		geospatial	spatial relations	location		generate maps		geospatial
quantity		part-to-whole	proportions		form and structure	see parts of whole, analyze text	composition	compositions (also of text)
association	correlate	correlation	relationships	hierarchy		relations between data points	relationship	correlations/ relationships

Typology of the Data Visualization Literacy Framework

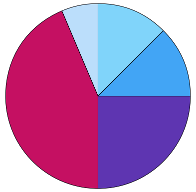
4

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

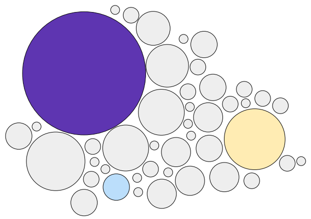
Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 30-31.

Visualization Types

Chart

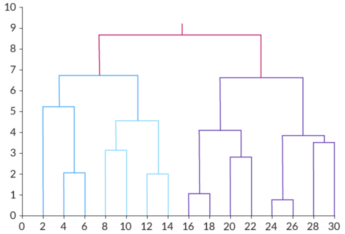


Pie Chart

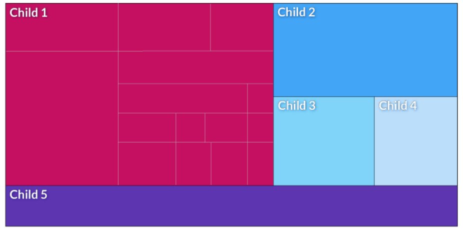


Bubble Chart

Tree

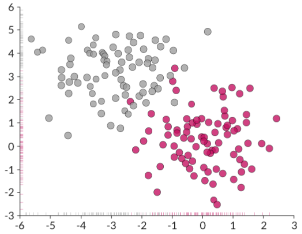


Dendrogram

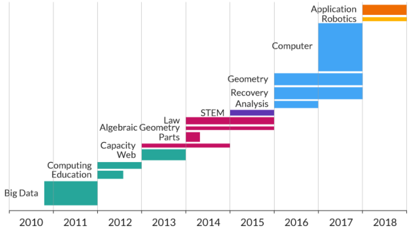


Tree Map

Graph

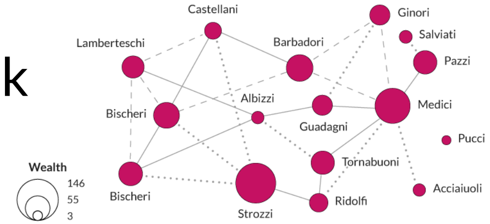


Scatter Graph

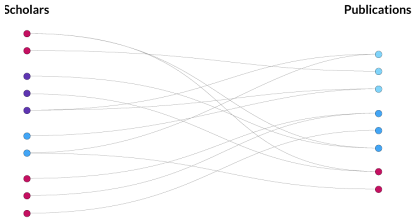


Temporal Bar Graph

Network

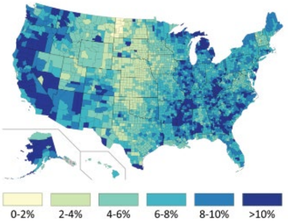


Force-Directed Network Layout

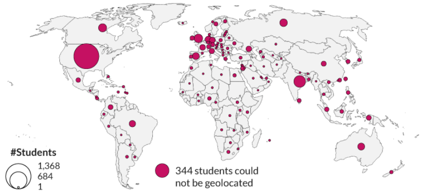


Bimodal Network Layout

Map



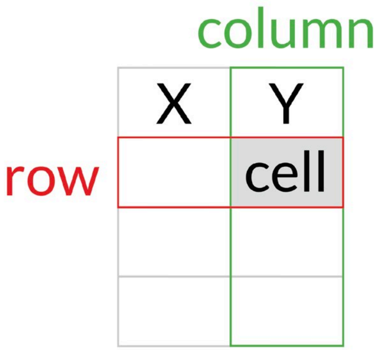
Choropleth Map



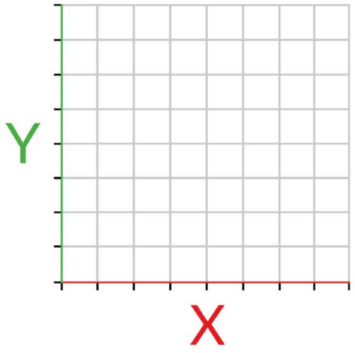
Proportional Symbol Map

Visualize: Reference Systems

Table
columns by rows

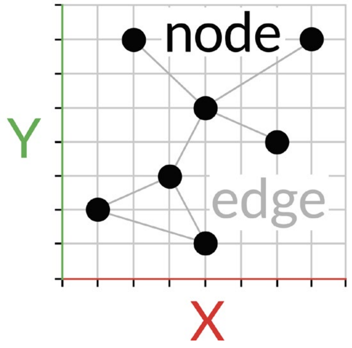


Graph
x-y coordinates



Map
latitude/
longitude

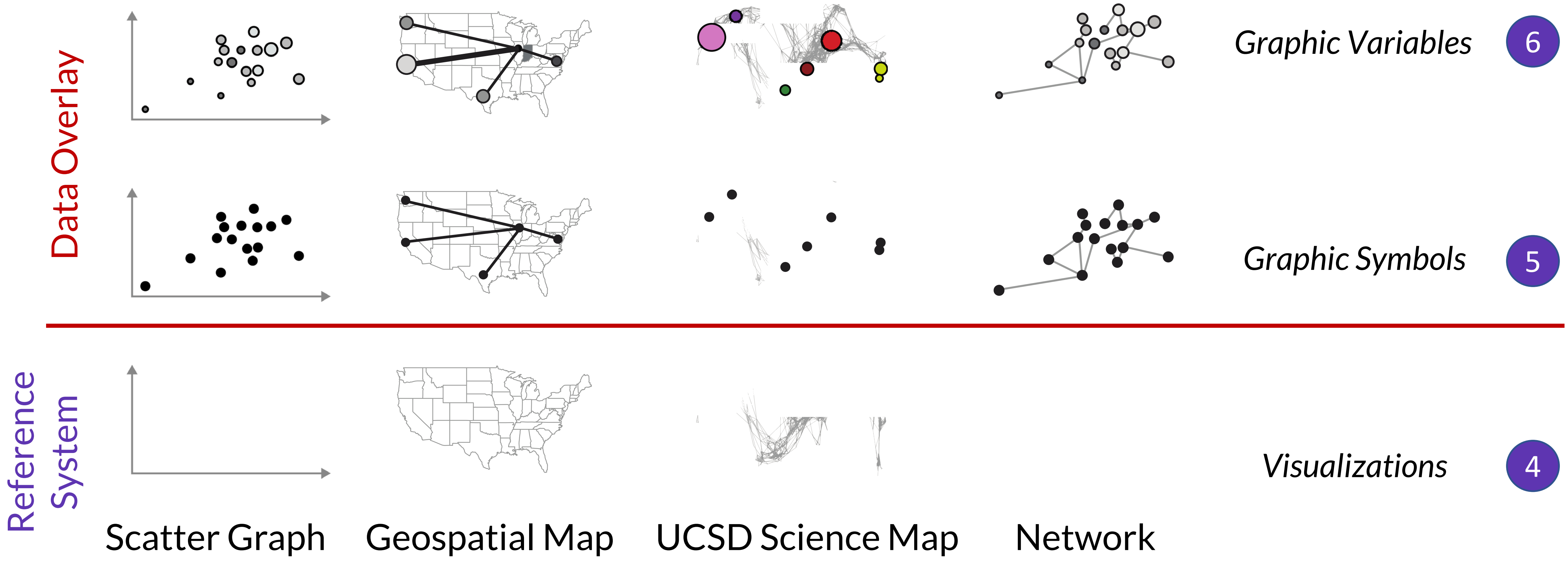
Network
local similarity



4

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

Visualize: Reference Systems, Graphic Symbols and Variables



Typology of the Data Visualization Literacy Framework

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

5

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

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

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

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

Typology of the Data Visualization Literacy Framework

6

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

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 34-35.

Graphic Variable Types

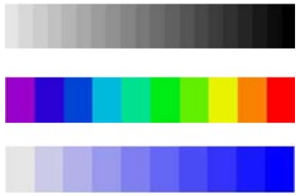
Position: x, y; possibly z

Form:

- Size
- Shape
- Rotation (Orientation)

Color:

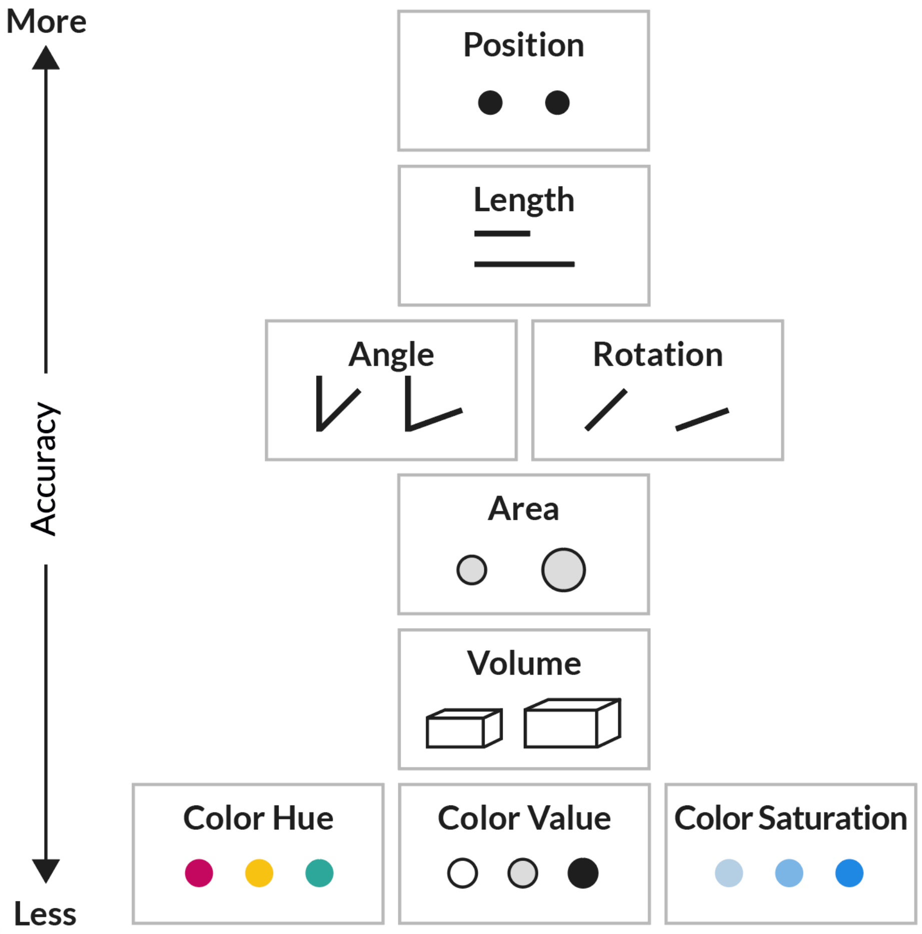
- Value (Lightness)
- Hue (Tint)
- Saturation (Intensity)



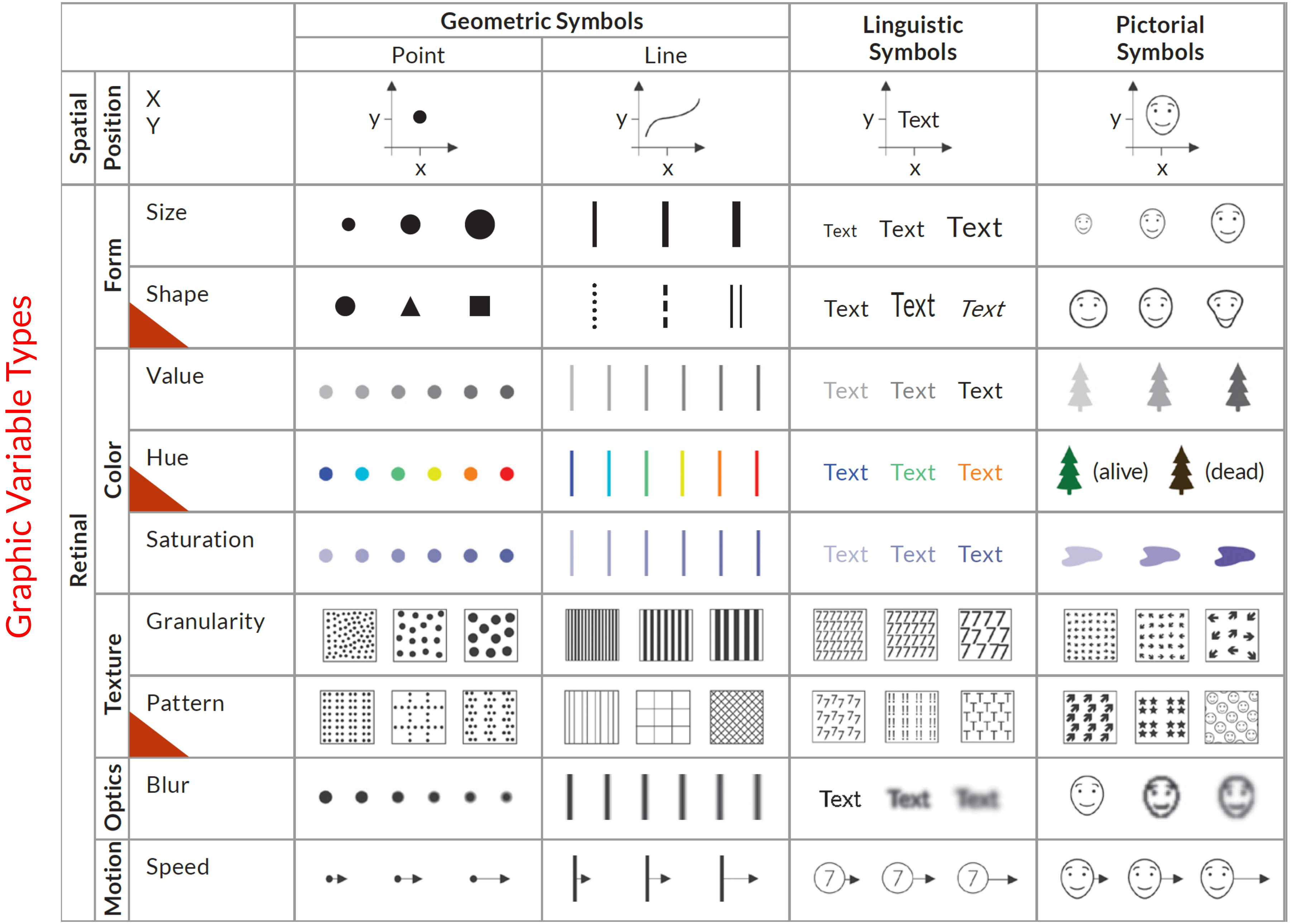
Optics: Blur, Transparency, Shading, Stereoscopic Depth

Texture: Spacing, Granularity, Pattern, Orientation, Gradient

Motion: Speed, Velocity, Rhythm



Graphic Symbol Types



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

Qualitative

Also called:

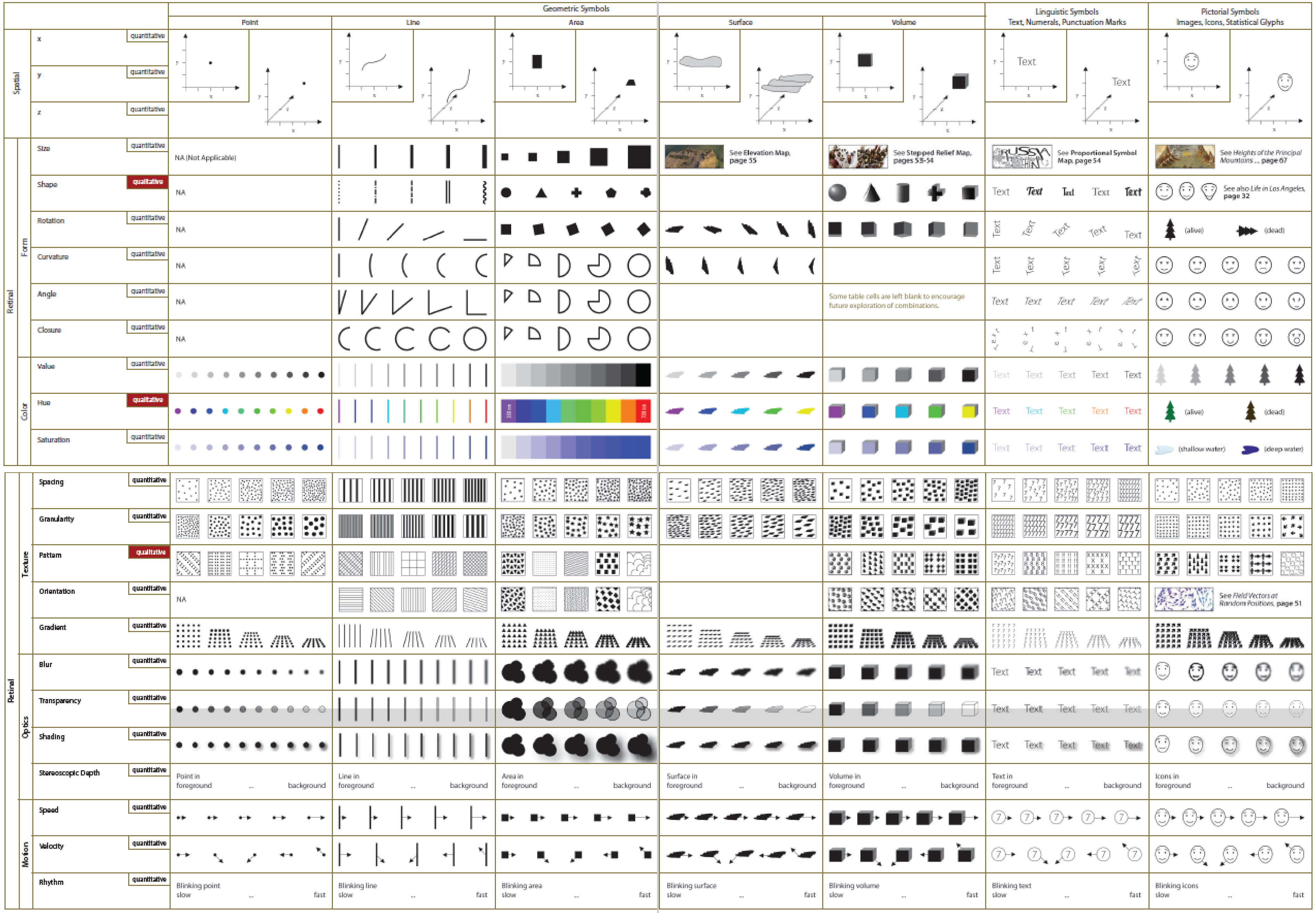
Categorical Attributes
Identity Channels

Quantitative

Also called:

Ordered Attributes
Magnitude Channels

Graphic Variable Types Versus Graphic Symbol Types



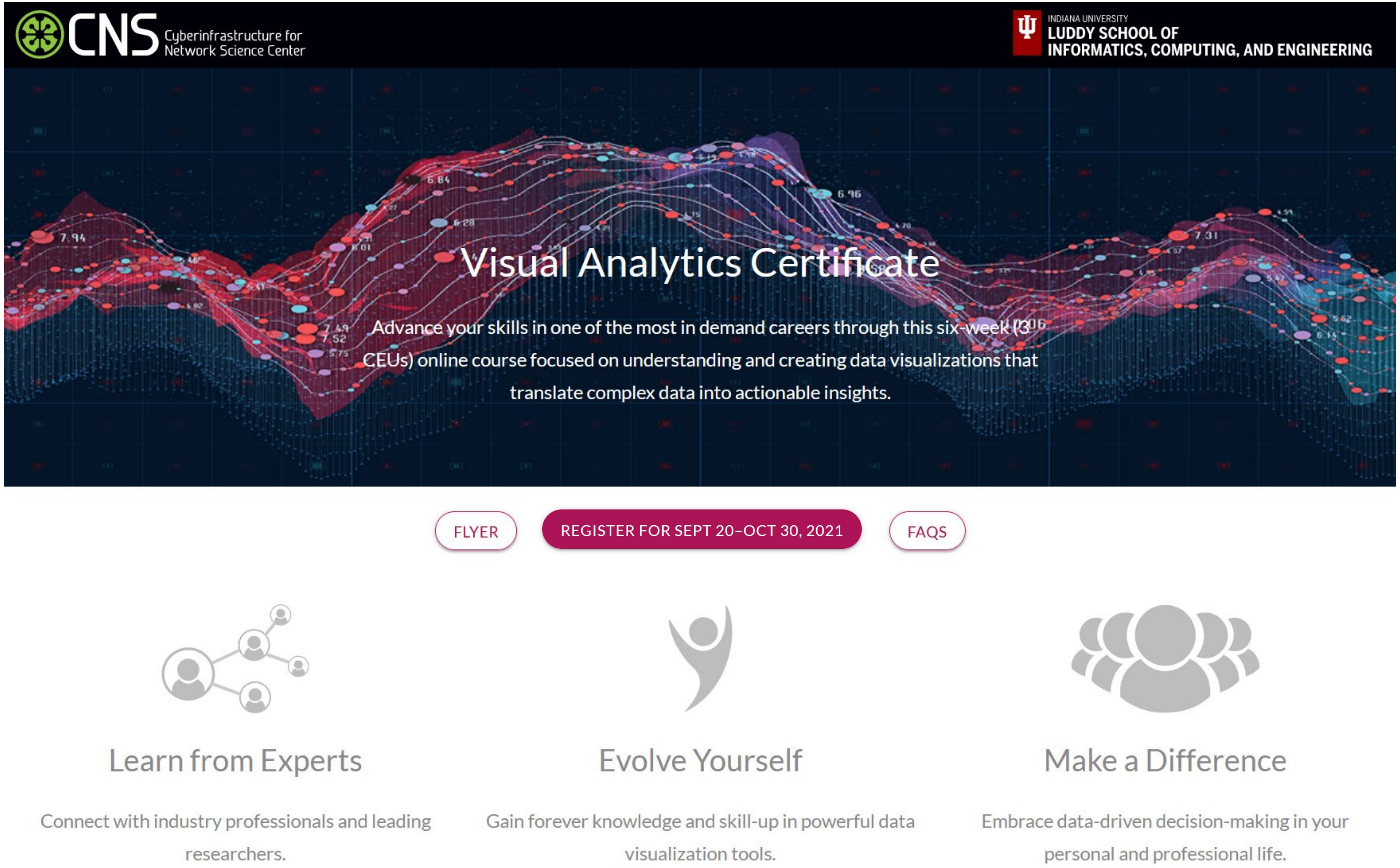
See *Atlas of Knowledge* pages 36-39 for complete table.

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Empower Yourself and Others! Data Visualization Literacy

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



https://visanalytics.cns.iu.edu

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

References

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255. http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1). http://www.pnas.org/content/vol101/suppl_1

Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. http://scimaps.org/atlas2

Q&A