Human Reference Atlas: Anatomical Structures, Cell Types & Biomarkers

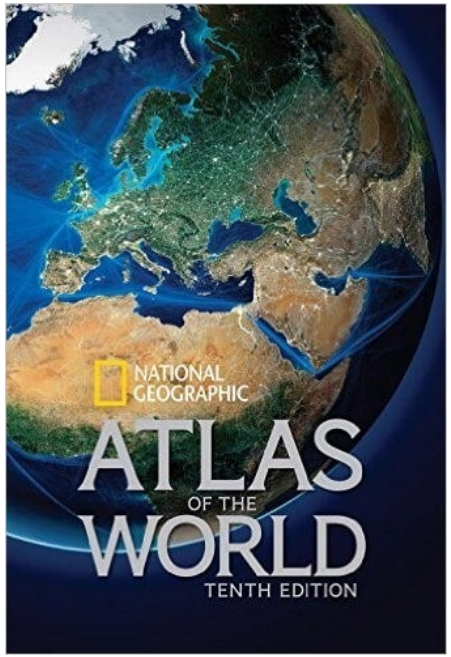
Katy Börner | @katycns
Victor H. Yngve Distinguished Professor of
Intelligent Systems Engineering & Information Science
Luddy School of Informatics, Computing, and Engineering
Indiana University, Bloomington, IN

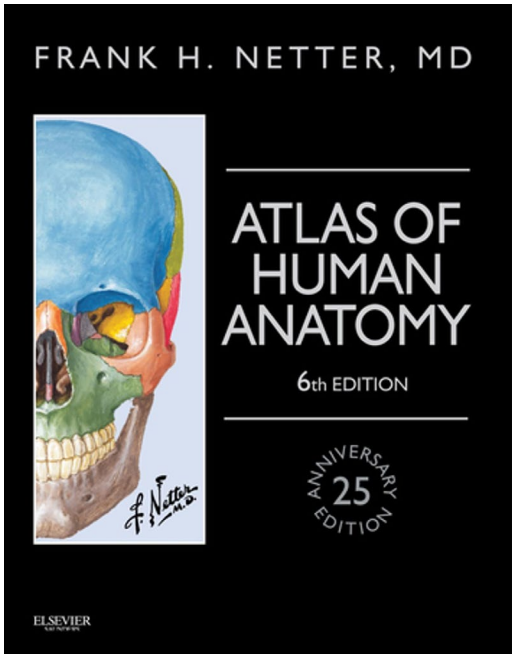
Seed Networks 2020 Annual Meeting
Virtual Event

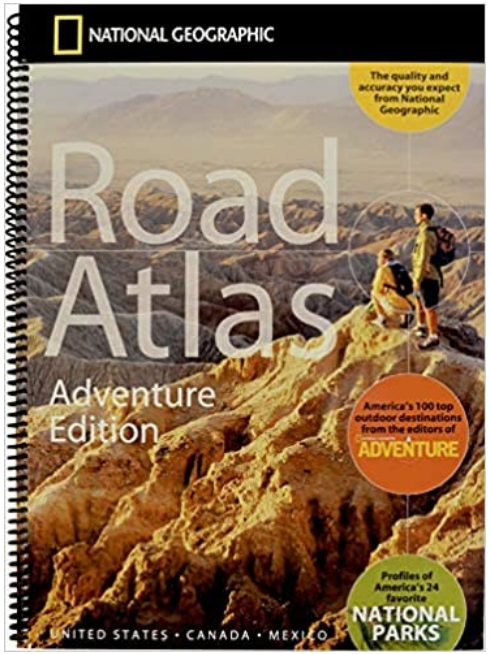
November 18, 2020

An **atlas** is an oversized, bound book of maps.

It has descriptive text, an index, possibly other data visualizations.

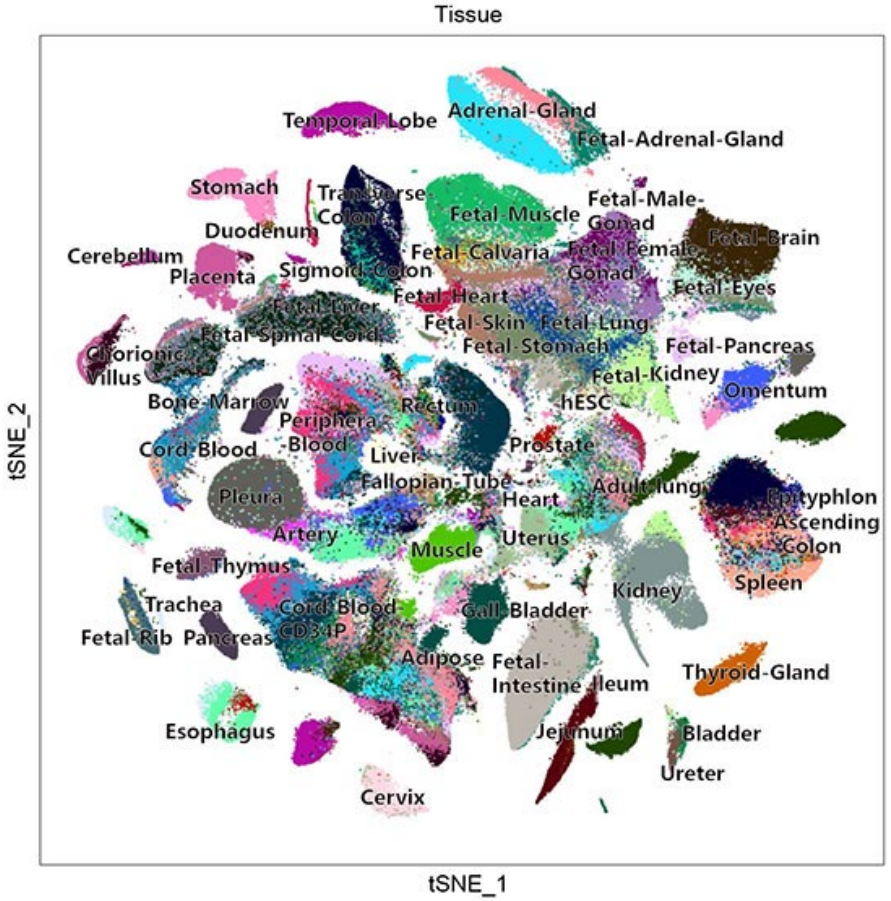






An human cell **atlas**
might show a landscape
of all cells, or

Maps of cells per tissue
type/anatomical structure.



Article | Published: 25 March 2020

Construction of a human cell landscape at single-cell level

Xiaoping Han, Ziming Zhou, [...] Guoji Guo

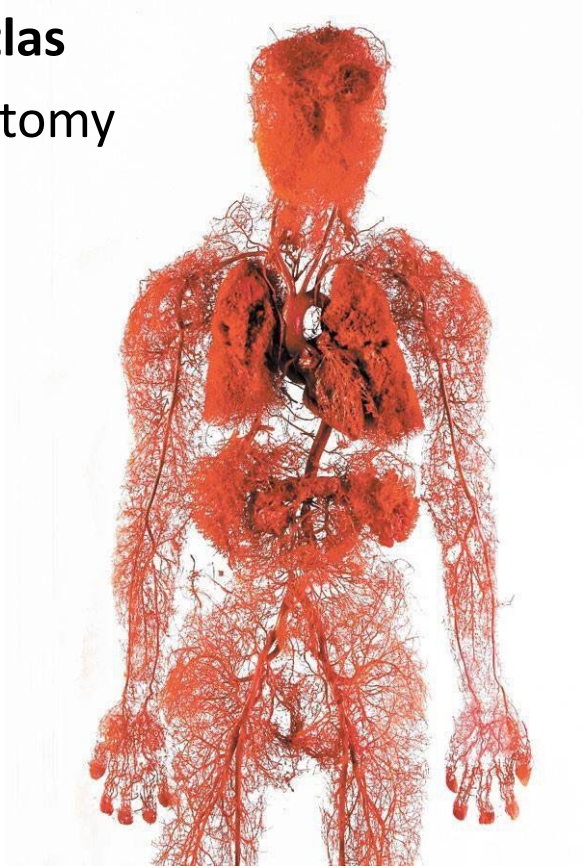
Nature **581**, 303–309(2020) | Cite this article

A human reference **atlas**

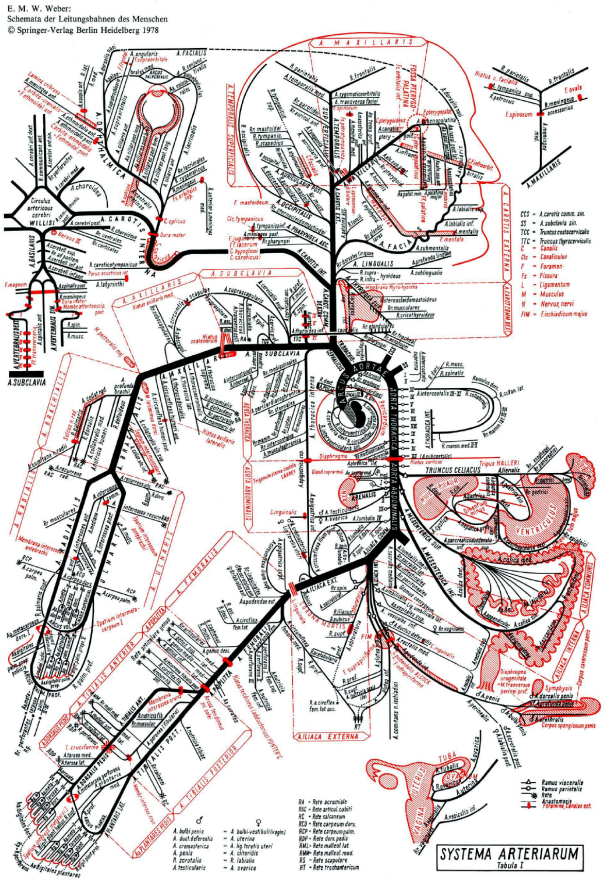
might use human anatomy

as a ‘basemap,’ or

an abstract space.



https://bodyworlds.com



Weber, 1978

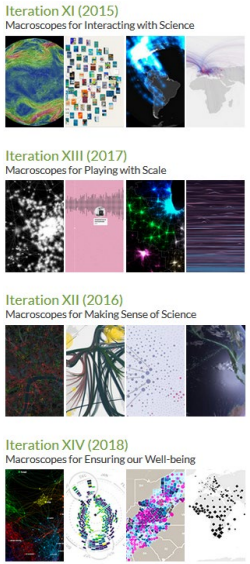
Places & Spaces: Mapping Science Exhibit

1st Decade (2005-2014)

Maps

2nd Decade (2015-2024)

Macroscopes



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

258
MAPMAKERS
from fields as disparate as art, urban planning, engineering, and the history of science.

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

20
MACROSCOPES
for touching all kinds of data.

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

354
PRESS ITEMS
including articles in *Nature*, *Science*, *USA Today*, and *Wired*.

http://scimaps.org

Acknowledgements

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 20 interactive macroscopes.

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

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

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

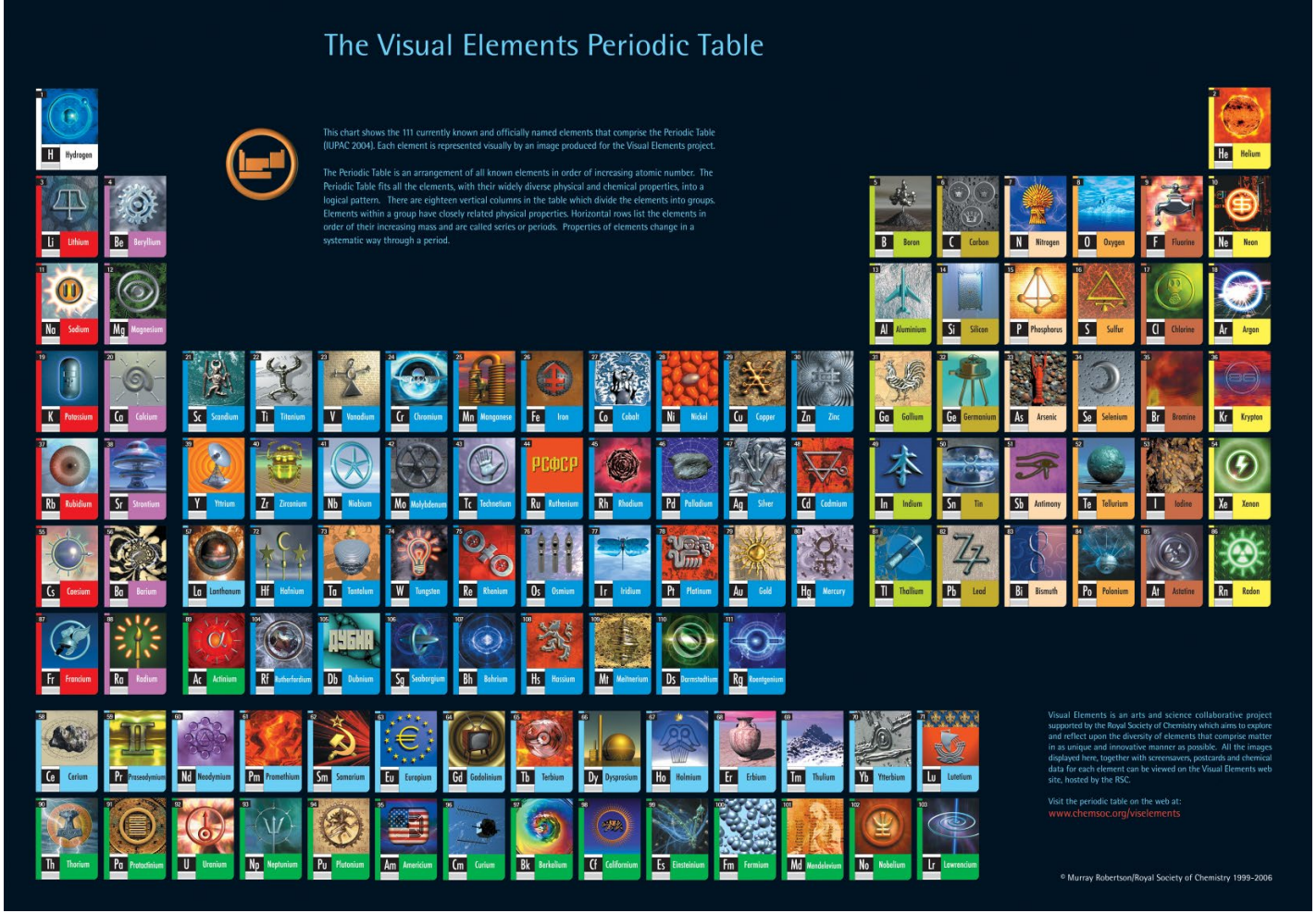
Stephen Uzzo
Vice President of Science and Technology for the New York Hall of Science

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

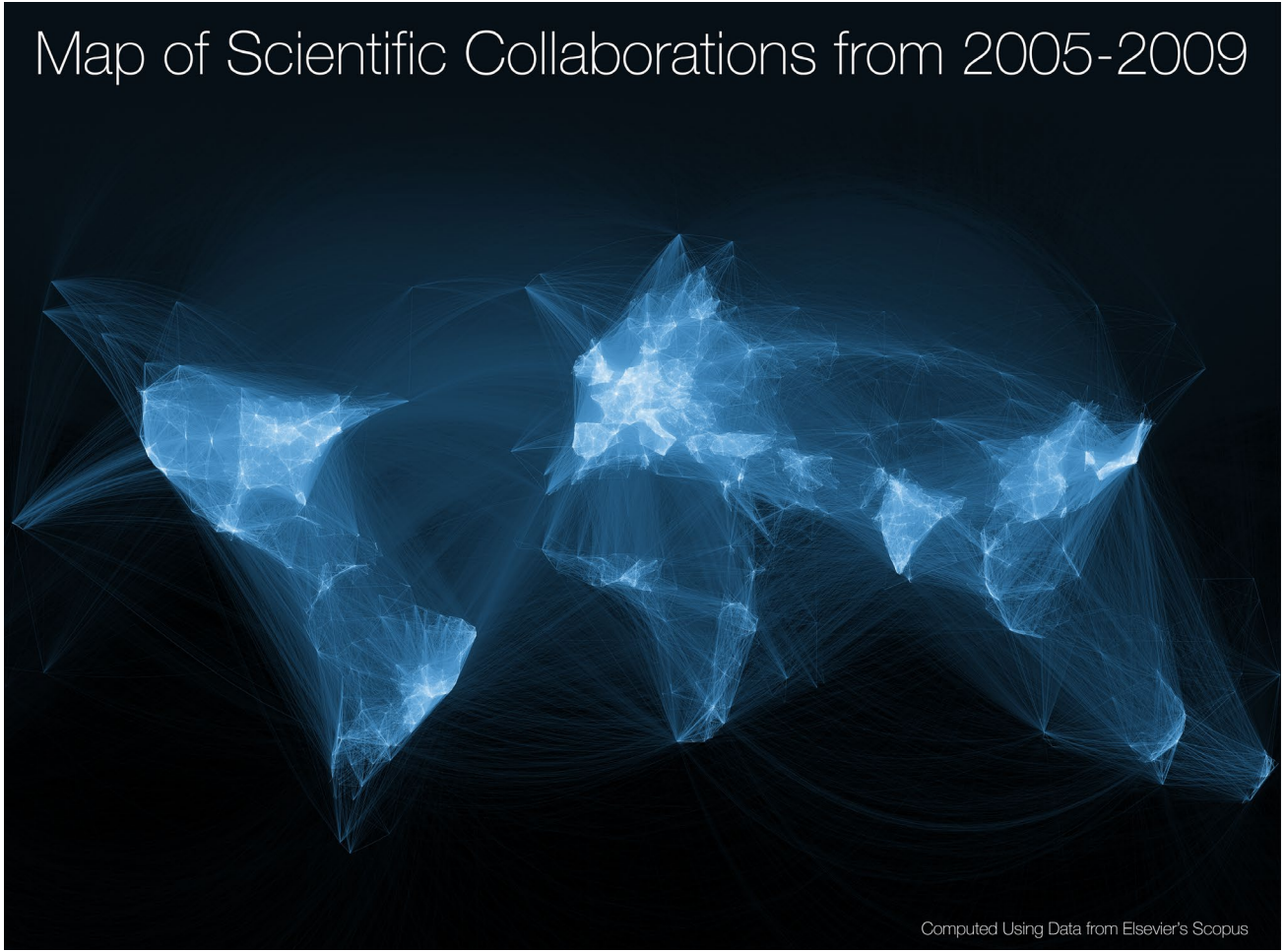
Benjamin Wiederkehr
Founding Partner and Managing Director of Interactive Things in Zürich, Switzerland

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

I.1 *Cosmographia World Map* – Claudius Ptolemy - 1482



II.2 *Visual Periodic Table of the Elements* - Murray Robertson and John Emsley - 2005



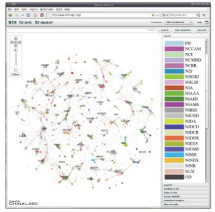
VII.6 Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012

A Topic Map of NIH Grants 2007

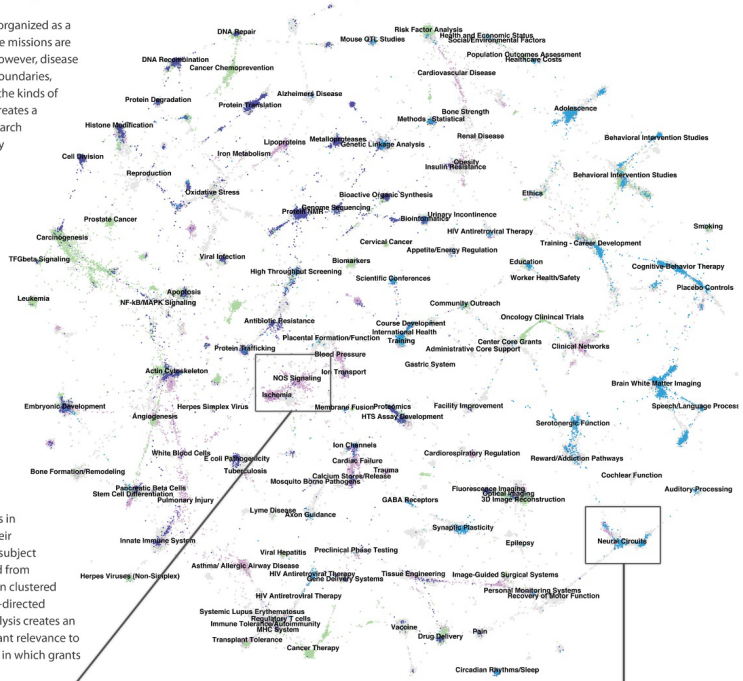
Bruce W. Herr II (Chalklabs & IU), Gully Burns (ISI), David Newman (UCI), Edmund Talley (NIH)

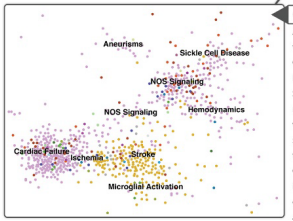
ChalkLabs IU IST UCIrvine

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.



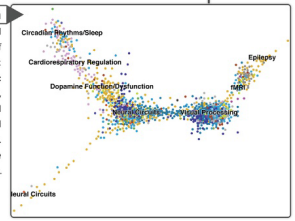


Cardiac Diseases Research

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)

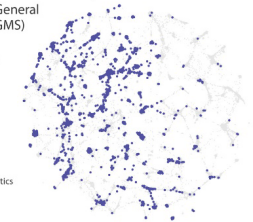
TOP 10 TOPICS

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

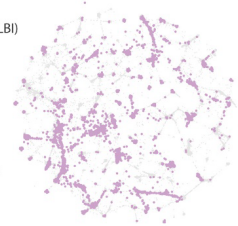
1. Bioactive Organic Synthesis
2. X-ray Crystallography
3. Protein NMR
4. Computational Models
5. Yeast Biology
6. Metalloproteases
7. Enzymatic Mechanisms
8. Protein Complexes
9. Invertebrate/Zebrafish Genetics
10. Cell Division



National Heart, Lung, and Blood Institute (NHLBI)

TOP 10 TOPICS

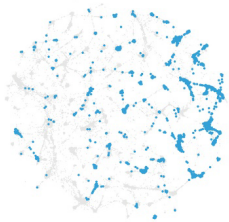
1. Cardiac Failure
2. Pulmonary Injury
3. Genetic Linkage Analysis
4. Cardiovascular Disease
5. Atherosclerosis
6. Hemostasis
7. Blood Pressure
8. Asthma/ Allergic Airway Disease
9. Gene Association
10. Lipoproteins



National Institute of Mental Health (NIMH)

TOP 10 TOPICS

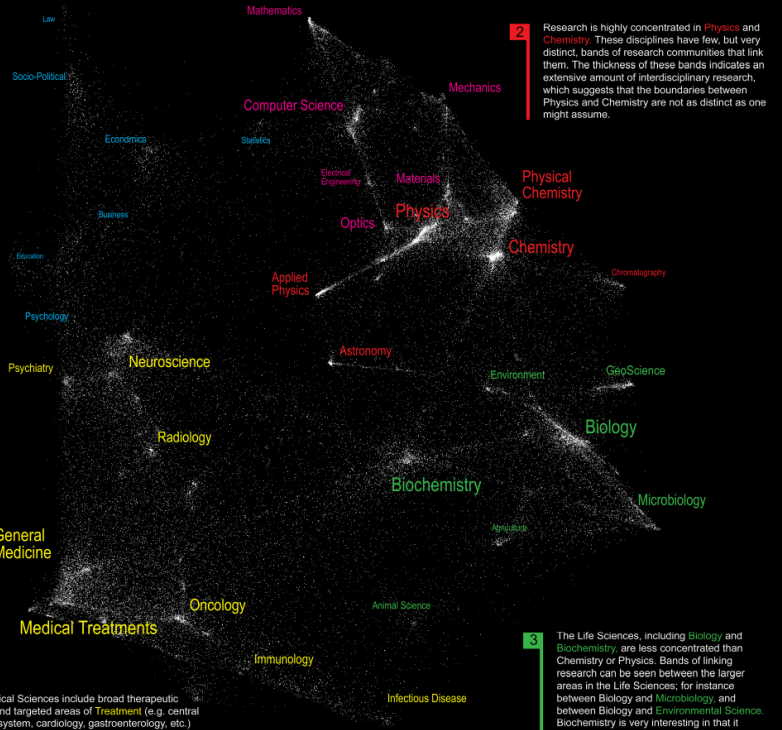
1. Mood Disorders
2. Schizophrenia
3. Behavioral Intervention Studies
4. Mental Health
5. Depression
6. Cognitive-Behavior Therapy
7. AIDS Prevention
8. Genetic Linkage Analysis
9. Adolescence
10. Childhood



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.

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.

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.

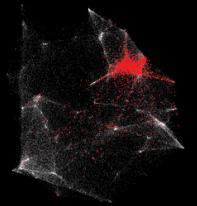
We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

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.

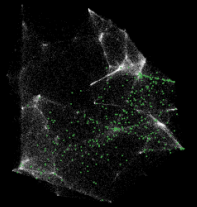
The map of science can be used as a tool for science strategy. This is the terrain in which organizations and institutions locate their scientific capabilities. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to explore,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



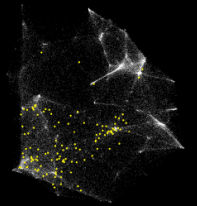
Nanotechnology

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

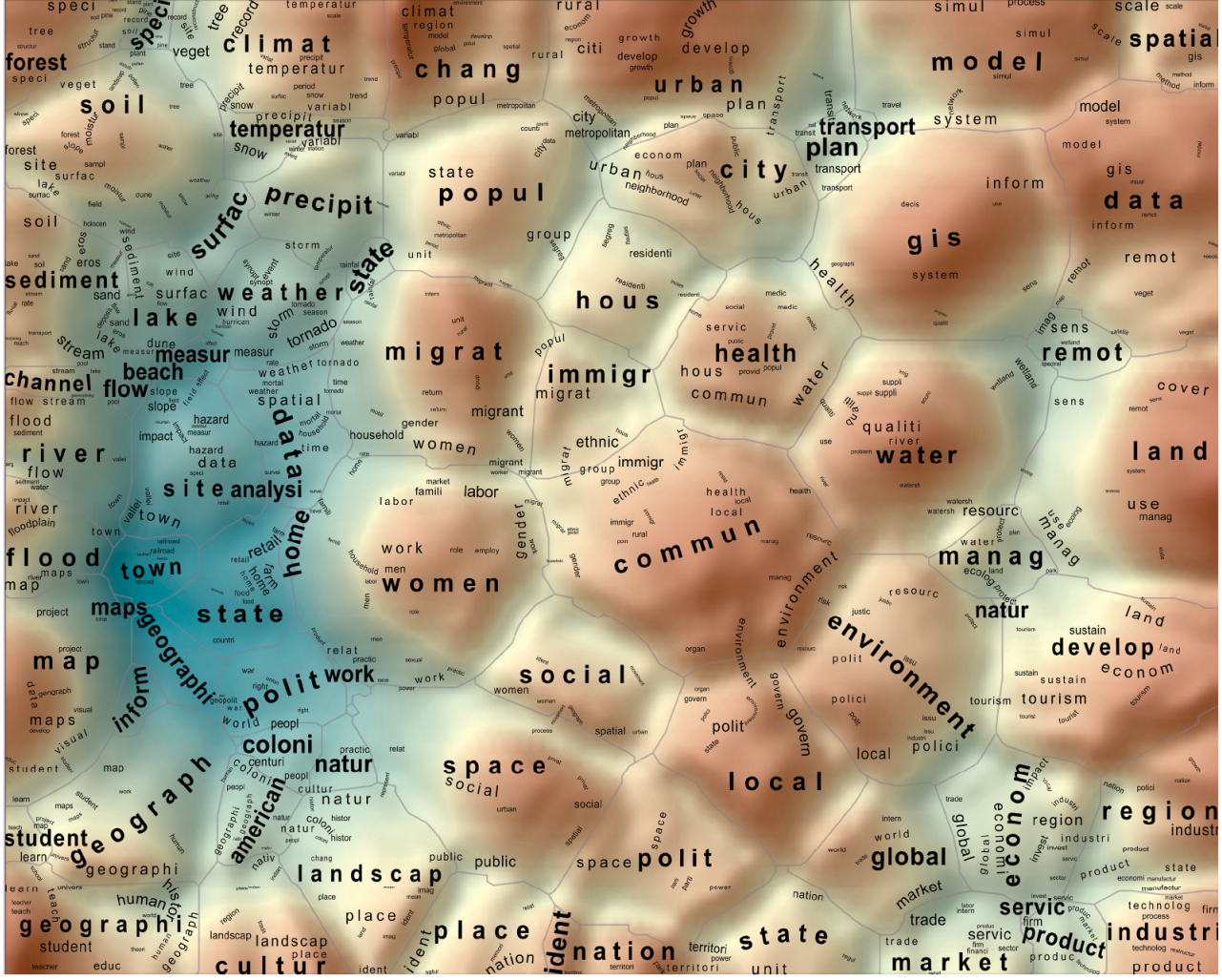
Research communities in proteomics are centered in Biochemistry. In addition, there is a heavy focus in the tools section of chemistry, such as Chromatography. The balance of the proteomics communities are widely dispersed among the Life and Medical Sciences.



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



I.9 *In Terms of Geography* – André Skupin - 2005

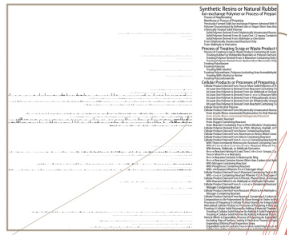
The US Patent Hierarchy

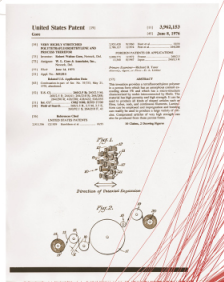
Impact

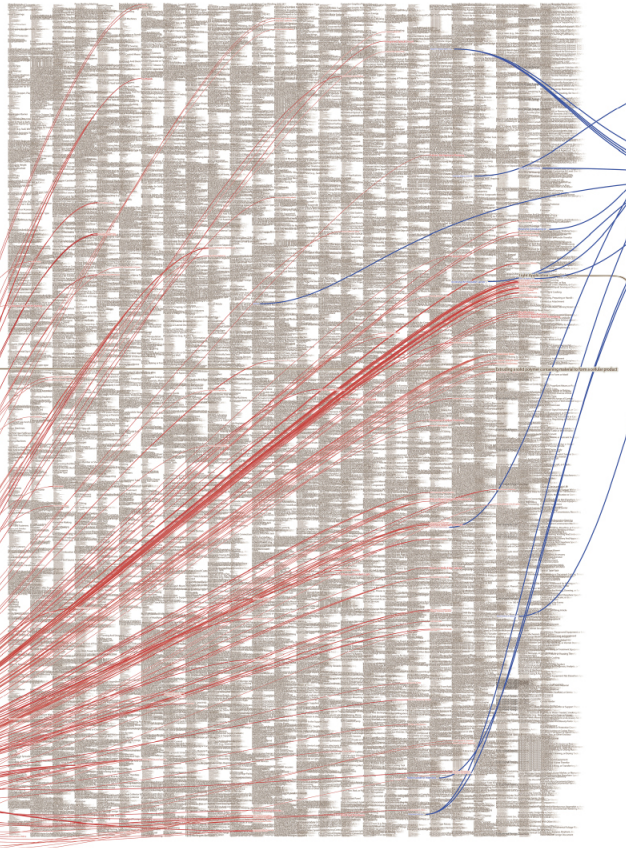
The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there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

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

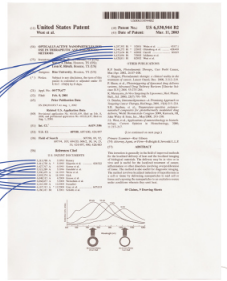
The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain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

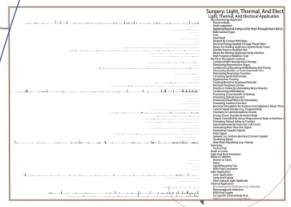






Prior Art





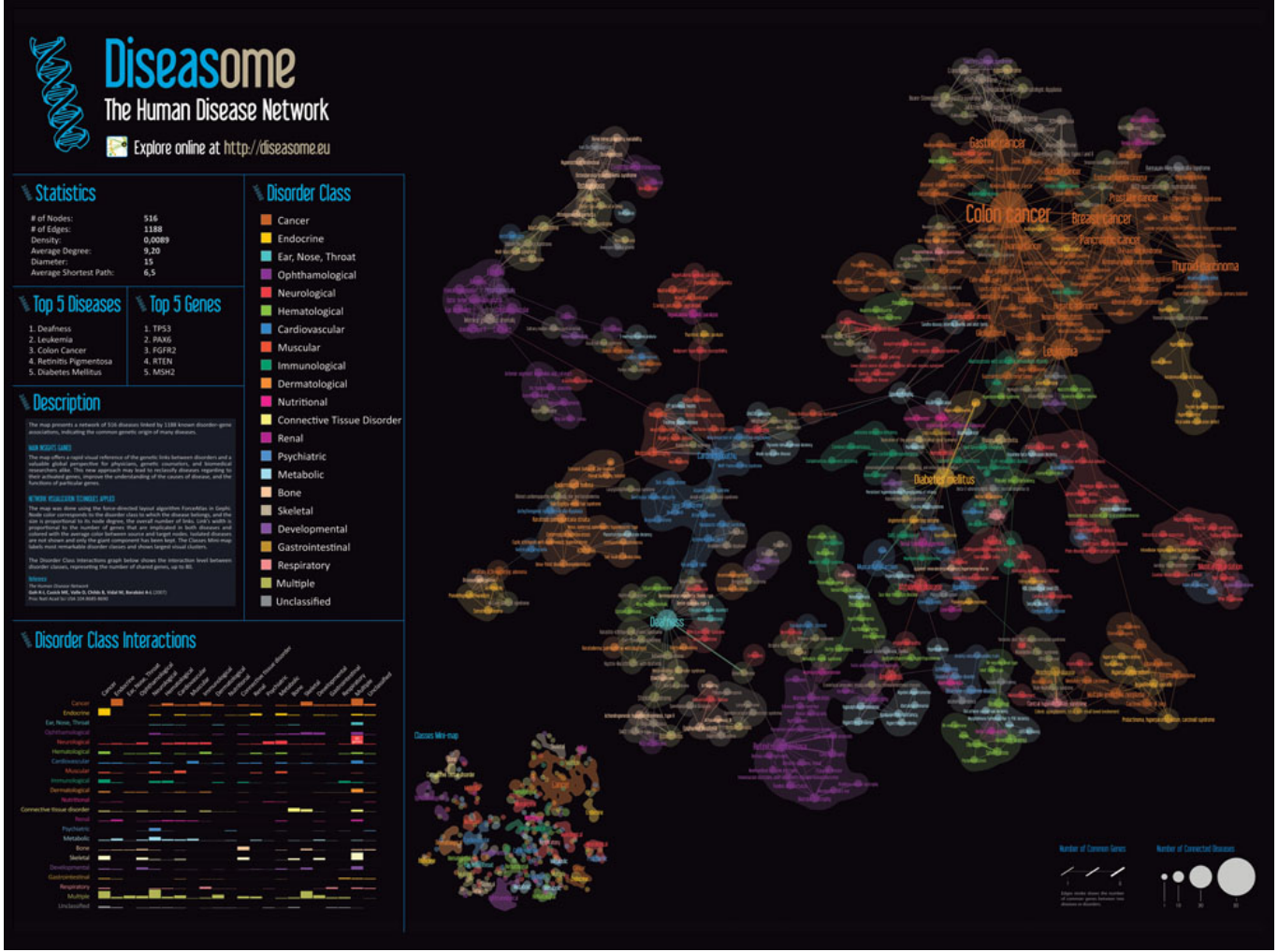
New patents often build on older ideas from many different categories. Here, blue lines originate in the sixteen categories that contain patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny gold spheres (with a diameter ten million times smaller than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

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

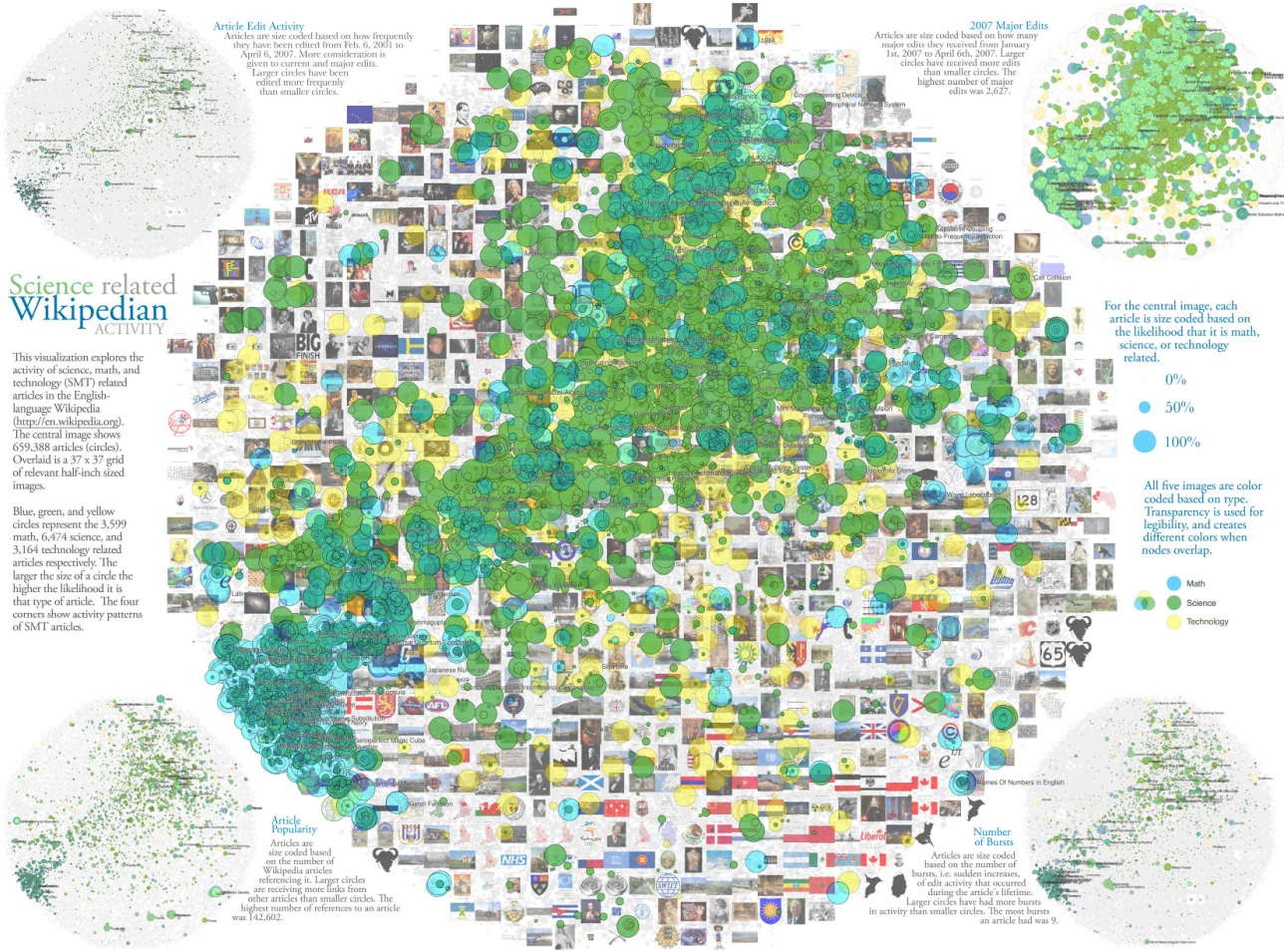
Something as simple as a bar chart helps people see how entities in a category relate to that category. Here, each bar encodes a "distance to prototype": how much each patent differs from an idealized "prototype patent" for that category. A measure like this can be based on statistics, computational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

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

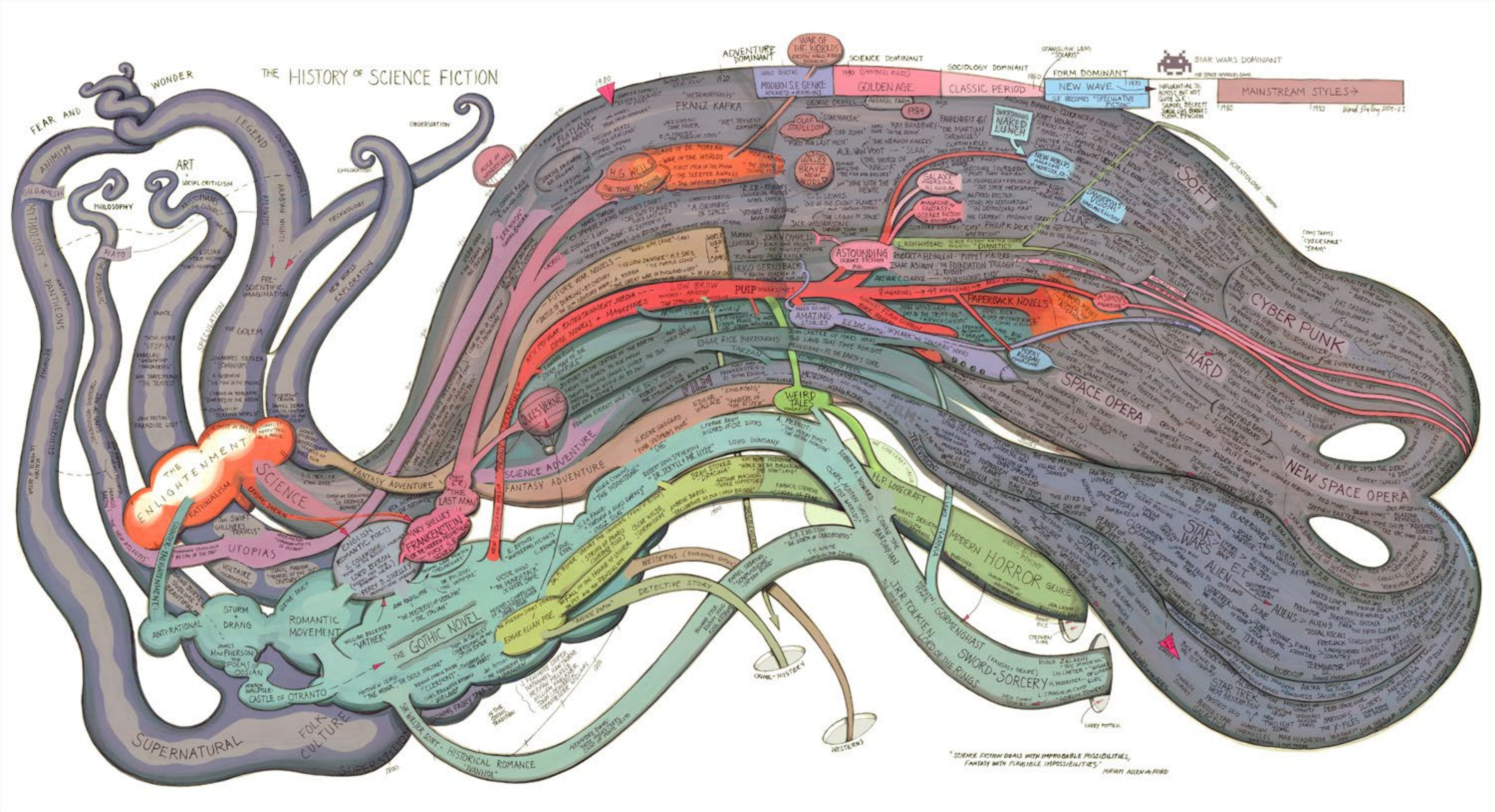
II.8 *Taxonomy Visualization of Patent Data* - Katy Börner, Elisha F. Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley - 2006



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

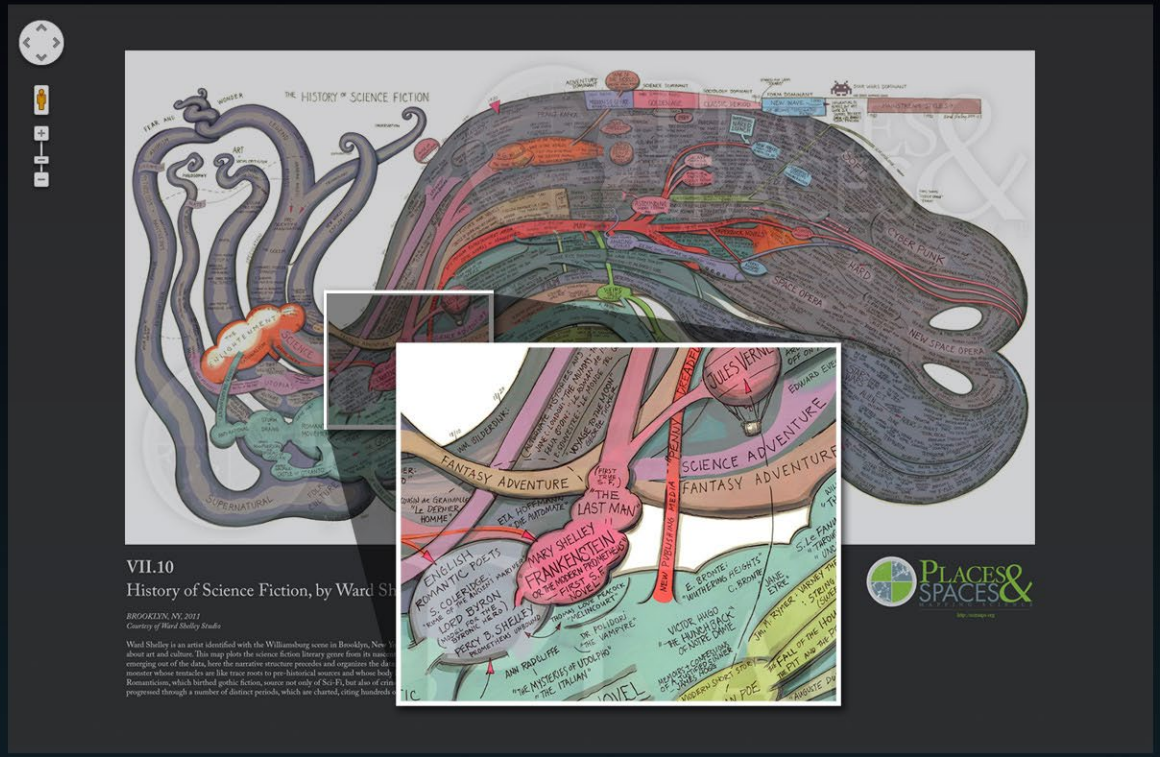


III.8 *Science-Related Wikipedian Activity* - Bruce W. Herr II, Todd M. Holloway, Elisha F. Hardy, Katy Börner, and Kevin Boyack - 2007



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

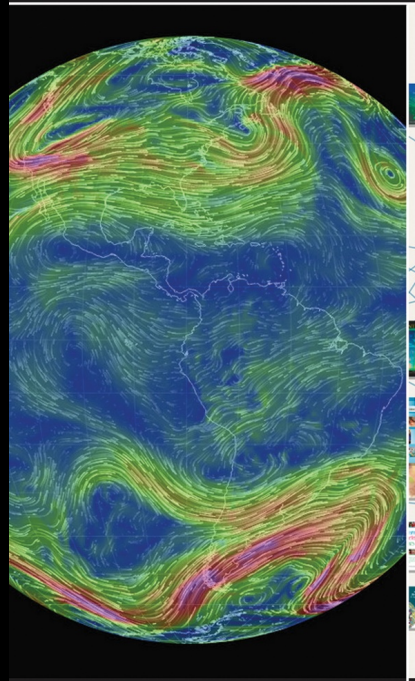
Check out our Zoom Maps online!



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

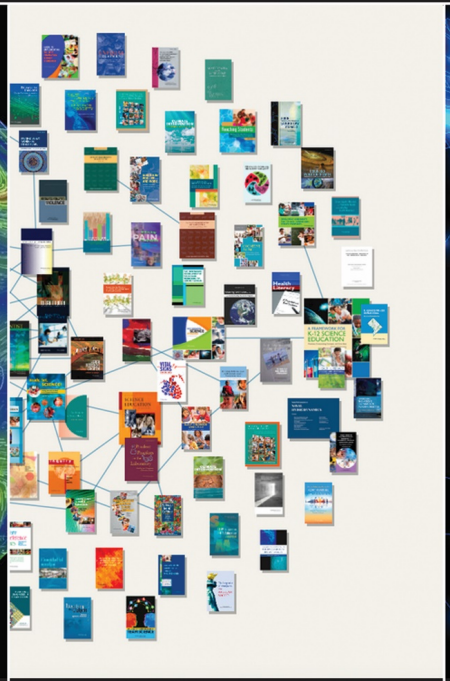
MACROSCOPES FOR INTERACTING WITH SCIENCE

PLACES & SPACES
MAPPING SCIENCE
scimaps.org



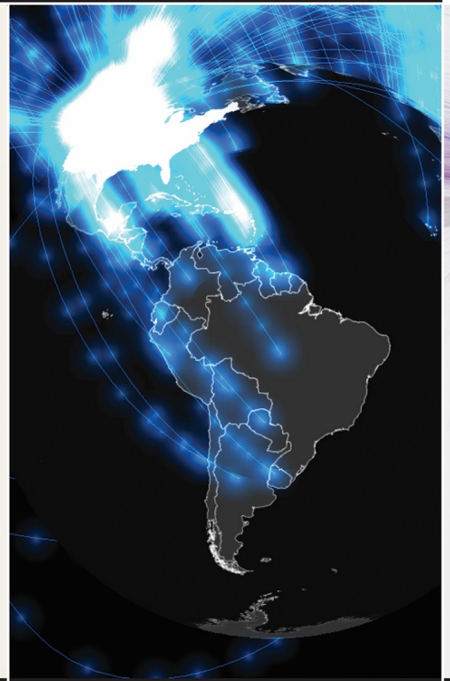
Earth

Weather on a worldwide scale



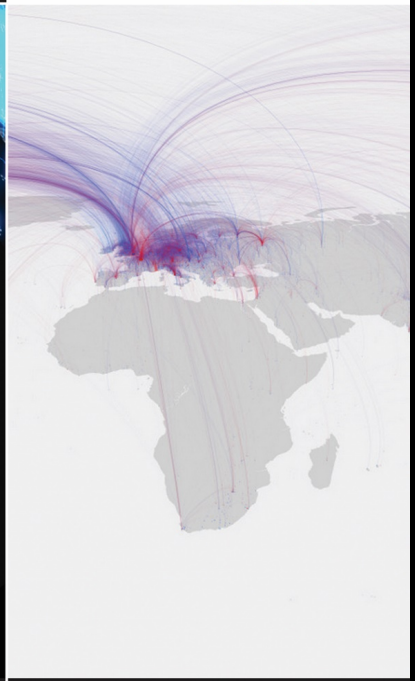
AcademyScope

Exploring the scientific landscape



Mapping Global Society

Local news from a global perspective



Charting Culture

2,600 years of human history in 5 minutes

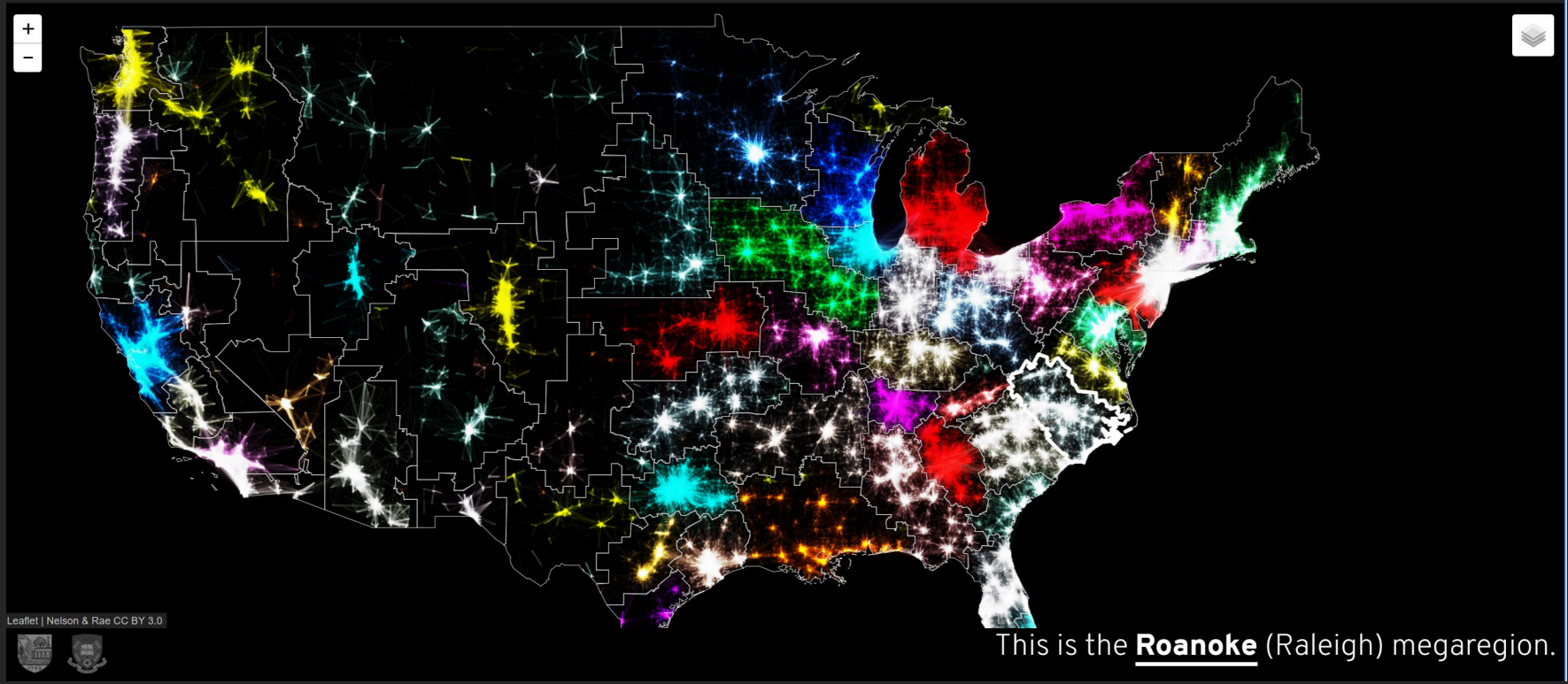
» Play with Scale » **Megaregions of the US**

MORE INFO

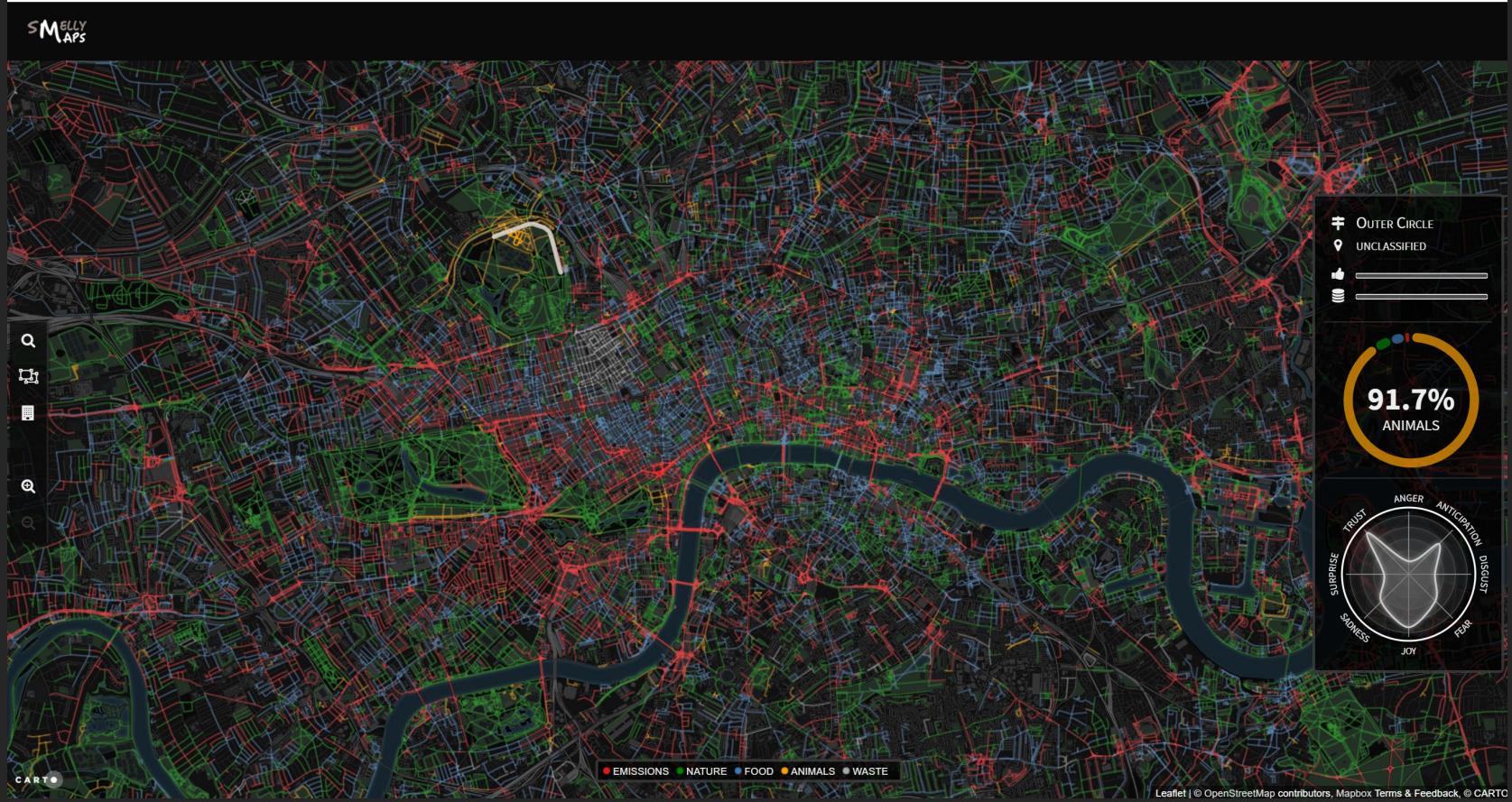
PLACES & SPACES MAPPING SCIENCE scimaps.org

THE MEGAREGIONS OF THE US

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



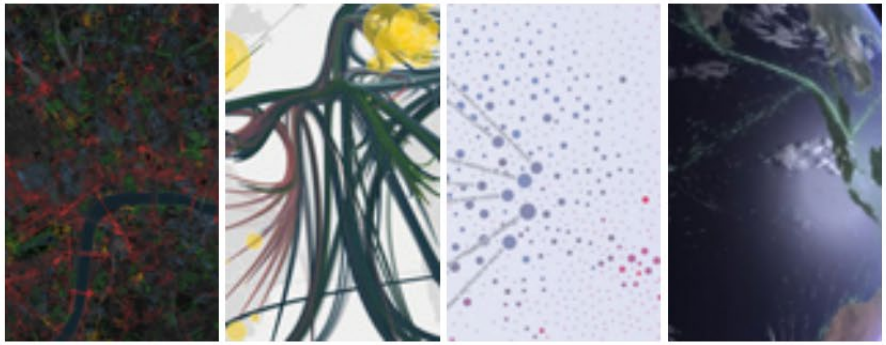
Megaregions of the US –Garrett Dash Nelson and Alasdair Rae – 2016



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

Iteration XII (2016)

Macroscopes for Making Sense of Science

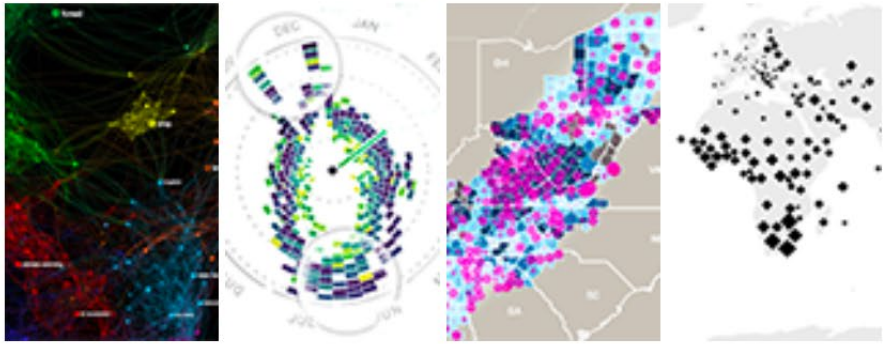


Iteration XIII (2017)

Macroscopes for Playing with Scale

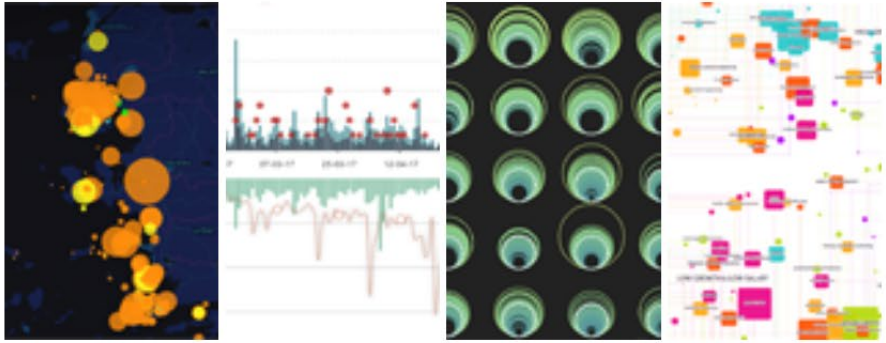
Iteration XIV (2018)

Macroscopes for Ensuring our Well-being



Iteration XV (2019)

Macroscopes for Tracking the Flow of Resources

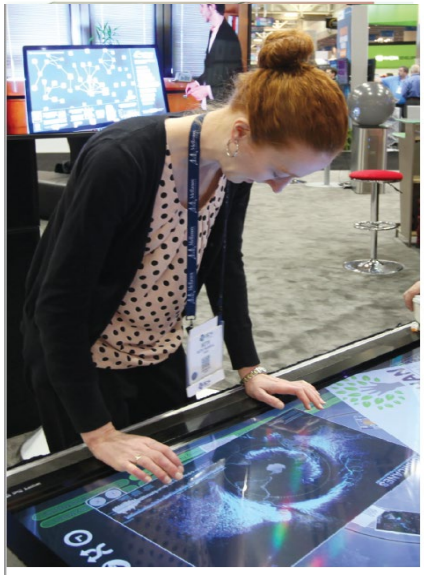


Geoffrey West, distinguished professor and past president, Santa Fe Institute, introduces Börner's Betazone talk at the World Economic Forum, Davos, Switzerland

"New Trends in eHumanities Research" workshop at the Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands

Ken Kennedy Institute for Information Technology, Rice University, Houston, TX

Illuminated Diagram display at the Smithsonian Folklife Festival, Washington, D.C.



Places & Spaces maps on a touch table at the International Conference for High Performance Computing, Networking, Storage, and Analysis, New Orleans, LA

Katy Börner debuts the exhibit at the University of Miami, Coral Gables, FL

100 science maps on display at the University of Miami, Coral Gables, FL

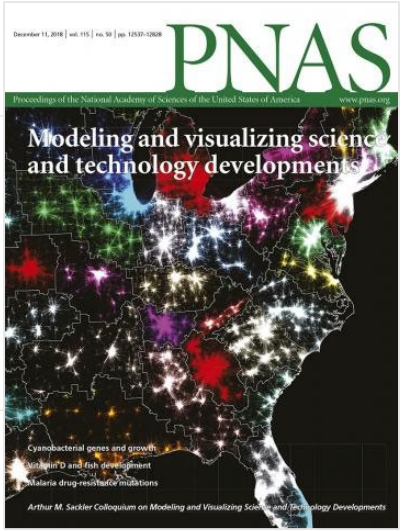
Maps on display at the European Commission, Directorate-General for Research and Innovation, Brussels, Belgium

Jax and the Big Data Beanstalk theater piece introduces visitors to data visualizations and science maps at the Science Museum of Minnesota, St. Paul, MN

Katy Börner presents "Maps & Macroscopes" at TEDxBloomington, Bloomington, IN

PNAS Proceedings of the National Academy of Sciences of the United States of America

Arthur M. Sackler Colloquium on Modeling and Visualizing Science and Technology Developments



Twin-Win Model: A human-centered approach to research success

Ben Shneiderman

PNAS December 11, 2018 115 (50) 12590-12594; first published December 10, 2018. https://doi.org/10.1073/pnas.1802918115

Forecasting innovations in science, technology, and education

FROM THE COVER

Katy Börner, William B. Rouse, Paul Trunfio, and H. Eugene Stanley

PNAS December 11, 2018 115 (50) 12573-12581; first published December 10, 2018. https://doi.org/10.1073/pnas.1818750115

How science and technology developments impact employment and education

Wendy Martinez

PNAS December 11, 2018 115 (50) 12624-12629; first published December 10, 2018. https://doi.org/10.1073/pnas.1803216115

Scientific prize network predicts who pushes the boundaries of science

Yifang Ma and Brian Uzzi

PNAS December 11, 2018 115 (50) 12608-12615; first published December 10, 2018. https://doi.org/10.1073/pnas.1800485115

The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms

C. Jara-Figueroa, Bogang Jun, Edward L. Glaeser, and Cesar A. Hidalgo

PNAS December 11, 2018 115 (50) 12646-12653; first published December 10, 2018. https://doi.org/10.1073/pnas.1800475115

https://www.pnas.org/modeling

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

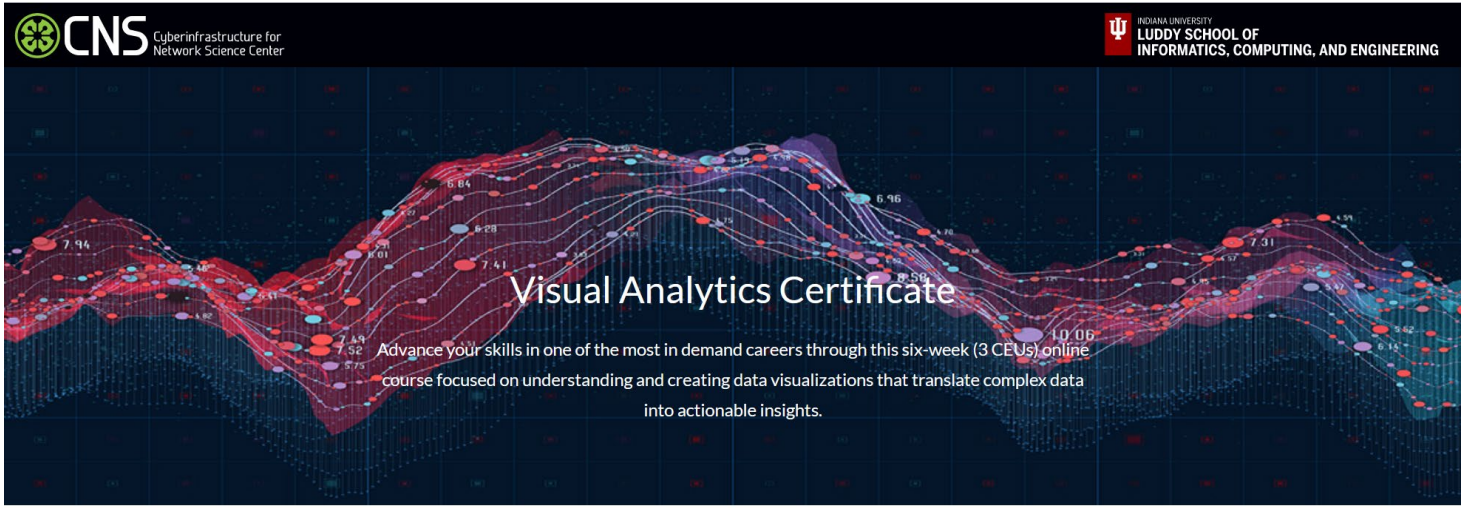
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

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

Börner, Katy (2021) **Atlas of Forecasts: Modeling and Mapping Desirable Futures**. The MIT Press.



FLYER | REGISTER FOR JAN 18–FEB 28, 2021 | FAQS

Learn from Experts

Connect with industry professionals and leading researchers.

Evolve Yourself

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

Make a Difference

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

https://visanalytics.cns.iu.edu

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

Human Reference Atlas: Anatomical Structures, Cell Types & Biomarkers

Acknowledgements

HuBMAP Consortium (https://hubmapconsortium.org)

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

NIH 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

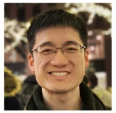
Andreas Bueckle
PhD Candidate

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

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

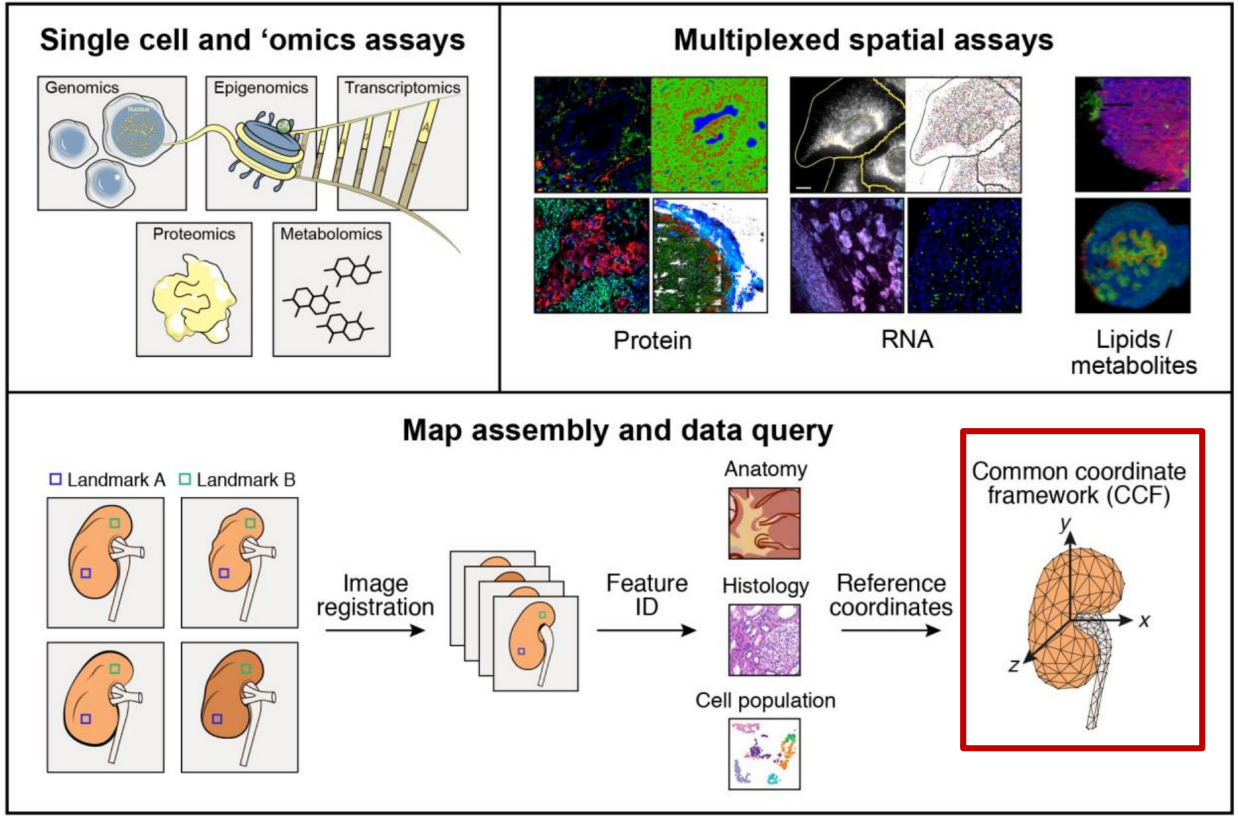


Fig. 3 | Map generation and assembly across cellular and spatial scales. HuBMAP aims to produce an atlas in which users can refer to a histological slide from a specific part of an organ and, in any given cell, understand its contents on multiple 'omic levels—genomic, epigenomic, transcriptomic, proteomic, 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.

Toward a Human Reference Atlas

Much recent research and ontology & reference organ design, including

- Rood, Jennifer E., Tim Stuart, Shila Ghazanfar, Tommaso Biancalani, Eyal Fisher, Andrew Butler, Anna Hupalowska, Leslie Gaffney, William Mauck, Gökcen Eraslan, John C. Marioni, Aviv Regev, and Rahul Satija. 2019. "Toward a Common Coordinate Framework for the Human Body." *Cell* 179 (7): 1455–1467. doi: 10.1016/j.cell.2019.11.019.
- Weber, Griffin M., Yingnan Ju, and Katy Börner. 2020. "Considerations for Using the Vasculature as a Coordinate System to Map All the Cells in the Human Body." *Frontiers in Cardiovascular Medicine* 7 (29). doi: 10.3389/fcvm.2020.00029.
- Allen Institute for Brain Science. 2020. "Allen Human Reference Atlas—3D, 2020." Version 1.0.0. Allen Brain Map Community Forum.
- Börner, Katy, Ellen M. Quardokus, Bruce W. Herr II, Leonard E. Cross, Elizabeth G. Record, Yingnan Ju, Andreas D. Bueckle, James P. Sluka, Jonathan C. Silverstein, Kristen M. Browne, Sanjay Jain, Clive H. Wasserfall, Marda L. Jorgensen, Jeffrey M. Spraggins, Nathan H. Patterson, Mark A. Musen, and Griffin M. Weber. 2020. "Construction and Usage of a Human Body Common Coordinate Framework Comprising Clinical, Semantic, and Spatial Ontologies." *arXiv*, July 28, 2020.

What is a CCF?

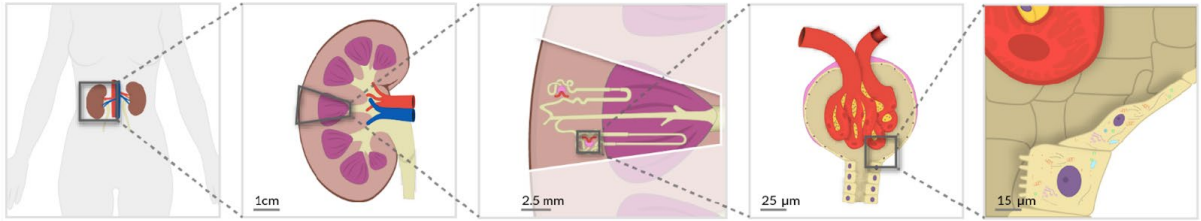
The Common Coordinate System (CCF) consists of ontologies and reference object libraries, computer software (e.g., user interfaces), and training materials that

- enable biomedical experts to semantically annotate tissue samples and to precisely describe their locations in the human body (“registration”),
- align multi-modal tissue data extracted from different individuals to a reference coordinate system (“mapping”) and,
- provide tools for searching and browsing HuBMAP data at multiple levels, from the whole body down to single cells (“exploration”).

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).

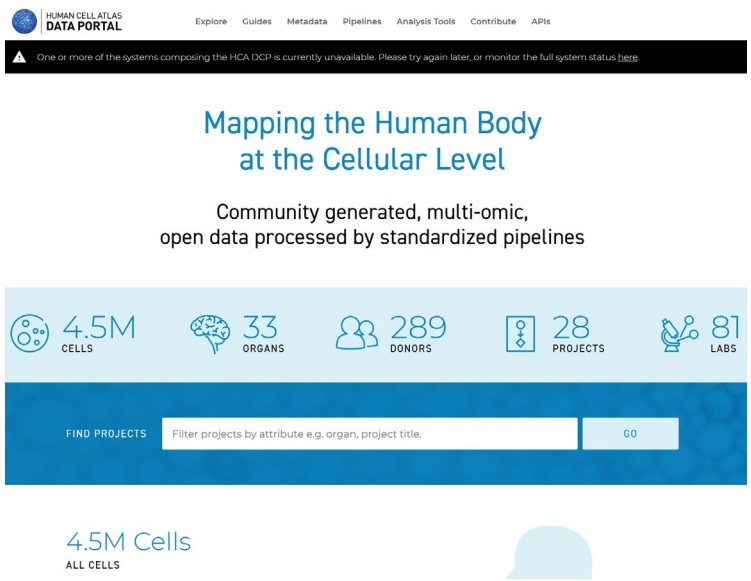


Body	Organ	Functional Tissue Unit	FTU Sub-structure(s)	Cellular
• Body • Kidney (Left, Right) • Aorta • Renal artery • Renal vein • Ureter	• Renal capsule • Renal pyramid • Renal cortex • Renal medulla • Renal calyx • Renal pelvis	• Nephron • Renal corpuscle • Proximal convoluted tubule • Loop of Henle • Distal convoluted tubule • Connecting tubule • Collecting duct	• Bowman's capsule • Glomerulus • Efferent arteriole • Afferent arteriole	• Parietal epithelial cell • Capillary endothelial cell • Mesangial cell • Podocyte

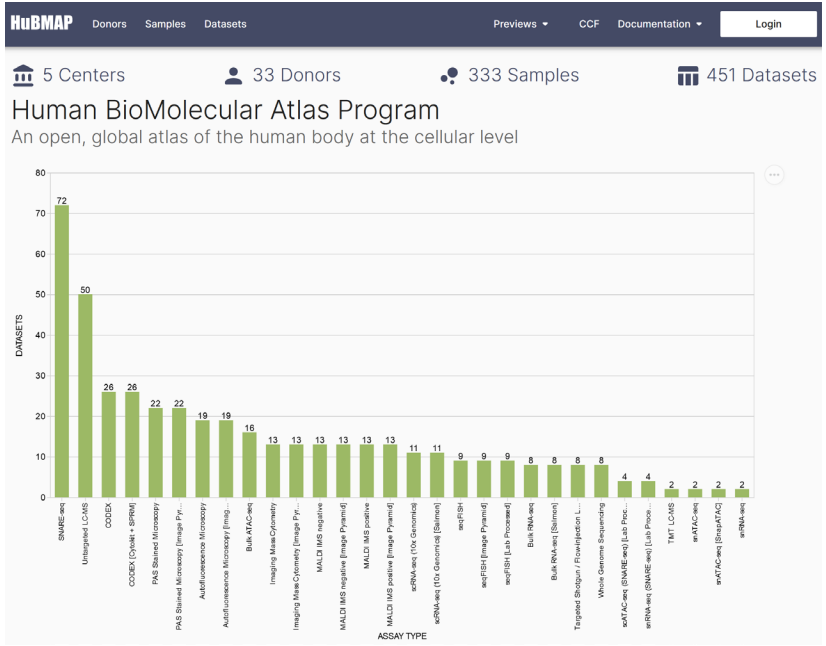
	HuBMAP	RBK	KPMP	SPARC	LungMAP	HTAN	HCA	GUDMAP	Gut Cell Atlas	BICCN	Allen Brain	TCGA	Wellcome	MRC	H2020	GTEx	Total
Kidney	1	1	1	0	0	0	1	1	0	0	0	1	1	1	0	1	9
Liver	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	3
Spleen	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	4
Heart	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	4
Lung	1	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	10
L intestine/Colon	1	0	0	1	0	1	1	0	1	0	0	1	0	0	0	1	7
S intestine	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
Bladder	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	5
Ureters	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2
Thymus	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2
Lymph nodes	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2
mediastinal lymph node	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Eye	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	3
Brain	0	0	0	0	0	0	1	0	0	1	1	1	0	0	1	1	6
Brain stem	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Cerebellum	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	3
Spinal cord	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	2
Pancreas	0	0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	5
Breast	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	1	5
Skin	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	3
Pediatric systems	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	2
Ovaries	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
Testes	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
Cervix	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Uterus	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	5
Blood	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	2
Bone	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Placenta	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Decidua	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Embryo	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
esophagus	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	3
hematopoietic system	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	2
immune system bulk	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Stomach	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	3
Thyroid	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
Prostate	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	3
Adrenal gland	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	3
Totals	**11**	**1**	**1**	**7**	**1**	**6**	**21**	**4**	**1**	**2**	**2**	**20**	**7**	**5**	**4**	**21**	**114**

Table compiled for, during, and after the NIH-HCA Joint Meeting in March 2020, https://hubmapconsortium.org/nihhca2020

Much data is becoming available, e.g.,



https://data.humancellatlas.org



https://portal.hubmapconsortium.org

ASCT+B Tables

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

Structure/Region	Substructure/Subregion	Cell Type	Subset of Marker Genes
Renal Corpuscle	Bowman's Capsule	Parietal epithelial cell	*CRB2*, CLDN1**
	Glomerulus	Podocyte	*NPHS2*, PODXL*, NPHS1**
		Capillary Endothelial Cell	*EHD3*, EMCN*, HECW2*, FLT1*, AQP1**
		Mesangial Cell	*POSTN*, PIEZO2*, ROBO1*, ITGA8**

Partial ASCT+B Table from

- El-Achkar et al. A Multimodal and Integrated Approach to Interrogate Human Kidney Biopsies with Rigor and Reproducibility: The Kidney Precision Medicine Project. bioRxiv. 2019, Updated Aug 2020. doi:10.1101/828665

Table 3: Cell types and associated markers from KPMP Pilot 1 transcriptomic studies. Asterisk denotes genes detected by more than one technology. *Italics*, genes detected by a single technology.

Structure/Region	Sub structure/Sub region	Cell Type	Abbreviation	Subset of Marker Genes	Pertinent negatives/comments
Renal Corpuscle	Bowman's Capsule	Parietal epithelial cell	PEC	*CRB2*, CLDN1**	
	Glomerulus	Podocyte	POD	*NPHS2*, PODXL*, NPHS1**	
		Capillary Endothelial Cell	GC-EC	*EHD3*, EMCN*, HECW2*, FLT1*, AQP1**	
		Mesangial Cell	MC	*POSTN*, PIEZO2*, ROBO1*, ITGA8**	
Tubules	Proximal Tubule	Proximal Tubule Epithelial Cell (general)	PT	*CUBN*, LRP2*, SLC13A1*, ALDOB*, GATM**	
		Proximal Convoluted Tubule Epithelial Cell Segment 1	PT-S1	*SLC5A2*, SLC5A12**	There is overlap among the segments
		Proximal Tubule Epithelial Cell Segment 2	PT-S2	*SLC22A6**	
		Proximal Tubule Cell Epithelial Segment 3	PT-S3	*PDZK1IP1*, MT1G**	
	Loop of Henle, Thin Limb	Descending Thin Limb Cell (general)	DTL	*CRYAB*, VCAM1*, AQP1*, SPP1**	*CLDN10* low
		Ascending Thin Limb Cell (general)	ATL	*CRYAB*, TACSTD2*, CLDN3**	*AQP1* low to none
	Loop of Henle, Thick Limb	Thick Ascending Limb Cell (general)	TAL	*SLC12A1*, UMOD**	*SLC12A3* low to none
		Cortex-TAL cell	C-TAL	*SLC12A1*, UMOD**	
		Medulla-TAL cell	M-TAL	*SLC12A1*, UMOD**	
		TAL-Macula Densa cell	TAL-MD	*NOS1*, SLC12A1**	
	Distal Convolution	Distal Convoluted Tubule Cell (general)	DCT	*SLC12A3*, TRPM6**	
		DCT type 1 cell	DCT-1	*SLC12A3*, TRPM6*	*SLC8A1, HSD11B2* (low to none)
		DCT type 2 cell	DCT-2	*SLC12A3*, SLC8A1*, HSD11B2*	Has CNT and DCT signature
	Connecting Tubule	Connecting Tubule Cell (general)	CNT	*SLC8A1*, CALB1, TRPV5*	*SLC12A3* low to none. IC or PC without *SLC8A1* could be in the CNT structure
		CNT-Principal Cell	CNT-PC	*SLC8A1*, AQP2*, SCNN1G**	
		CNT-Intercalated Cell	CNT-IC	*SLC8A1*, CA2, ATP6VOD2**	
		CNT-IC-A cell	CNT-IC-A	*SLC8A1*, SLC4A1*, SLC26A7**	
		CNT-IC-B cell	CNT-IC-B	*SLC8A1*, SLC26A4*, SLC4A9**	
	Collecting Duct	Collecting duct (general) cell	CD	*GATA3**	*GATA3* may be in subpopulation of DCT, CNT and vSMC/P. SLC8A1, CALB1, TRPV5
		CD-PC (general)	CD-PC	*AQP2*, AQP3*, FXYD4*, SCNN1G*, GATA3**	
		C-CD-PC	C-CD-PC		
		M-CD-PC	M-CD-PC		
		Outer medulla-CD-PC	OM-CD-PC		
		Inner Medulla-CD cell	IM-CD	*AQP2*, SLC14A2*	
		Transitional PC-IC cell	tPC-IC	*FXYD4*, SLC4A9*/SLC26A7**	(low to none); Low to No *CALCA* and *KIT* in C-CD-IC-A. It may not be possible to assign IC or PC to CNT or CD structures without regional information of their source.
		CD-IC (general) cell	CD-IC	*CA2, ATP6VOD2**	
		CD-IC-A (general) cell	CD-IC-A	*SLC4A1, SLC26A7*, TMEM213**	
		C-CD-IC-A cell	C-CD-IC-A	*SLC26A7*, SLC4A1**	
		M-CD-IC-A cell	M-CD-IC-A	*SLC26A7*, SLC4A1, KIT*, CALCA*	
		CD-IC-B (general) cell	CD-IC-B	*SLC4A9*, SLC26A4**	
		C-CD-IC-B cell	C-CD-IC-B		
		M-CD-IC-B cell	M-CD-IC-B		
Vessels	Endothelial Cells (non-glomerular)	Endothelial Cell (general)	EC	*EMCN*, PECAM1*, FLT1**	
		EC-Afferent/Efferent Arteriole	EC-AEA	*SERPINE2*, TM4SF1**	likely *PALMD*
		EC-Peritubular capillaries	EC-PTC	*PLVAP**	
		EC-Descending Vasa Recta	EC-DVR	*TM4SF1*, PALMD*	
		EC-Ascending Vasa Recta	EC-AVR	*DNASEIL3**	low to none
		EC-Lymphatics	EC-LYM	*MMRN1*, PROX1*	
Structure/Region	**Sub structure/Sub region**	**Cell Type**	**Abbreviation**	**Subset of Marker Genes**	**Pertinent negatives/comments**
Interstitium	Stroma (non-glomerular)	Vascular Smooth Muscle/Pericyte (general)	vSMC/P	*TAGLN*, ACTA2*, MYH11*, NTRK3, MCAM*	
		vSMC/P-Renin	vSMC/P-REN	*REN*	
		Fibroblast	FIB	*DCN*, ZEB2, C7, LUM*	
	Immune	Macrophages-Resident	MAC-R	*CD163*, IL7R**	
		Macrophage	MAC	*S100A9*	
		Natural Killer Cell	NKC	*NKG7*	
		Dendritic Cell	DC	*APOE*	
		Monocyte	MON	*C1QA, HLA-DRA*	
		T lymphocyte (general)	T	*CD3*	
		T Cytotoxic	T-CYT	*GZMA*	
		B lymphocyte	B	*IGJ*	

El-Achkar et al. A Multimodal and Integrated Approach to Interrogate Human Kidney Biopsies with Rigor and Reproducibility: The Kidney Precision Medicine Project. bioRxiv. 2019, Updated Aug 2020. doi:10.1101/828665

ASCT+B Table Working Group

Lead by Katy Börner and Jim Gee; Ellen M Quardokus serves as Knowledge Manager

Meetings take place monthly to review and approve tables, formalize and unify table design language, discuss and expand table usage, see WG Charter.

Next meetings: Dec 3, 1:30p EST. **In 2021:** Jan 6, Feb 3, March 3, 11a-noon ET.
Please register to receive invites and updates.

SOP for ASCT+B Tables

ASCT+B for 10 organs on 9/14/2020, 9:45am:

Organ Name	#AS	#CT	#B	#AS-CT	#CT-B
Brain	21	127	254	127	346
Heart	23	16	35	73	42
Kidney	39	53	83	55	135
Large Intestine	22	33	45	306	72
Liver	16	27	34	29	35
Lung	18	62	103	110	128
Lymph Nodes	34	30	50	63	110
Skin	14	32	57	37	99
Small intestine	20	32	48	196	57
Spleen	33	26	46	48	72

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

SOP for Construction, Review, Revision of Anatomical Structure and Cell Types and Biomarker (ASCT+B) Tables

Authors: Ellen M. Quardokus, Lisel Record, Bruce W. Herr II, Hrishikesh Paul, Katy Börner

September 18, 2020

Introduction

Anatomical Structures, Cell Types, plus Biomarkers (ASCT+B) tables aim to capture the nested *part_of* structure of anatomical human body parts, the typology of cells, and biomarkers used to identify cell types (e.g., gene, protein, lipid or metabolic markers). The tables are authored and reviewed by an international team of anatomists, pathologists, physicians, and other experts.

Identification of Subject Matter Experts (SMEs)

- CCF Experts (cross-consortium team lead by MC-IU) invite leading organ experts to contribute to the design of ASCT+B tables.
- Leading organ experts submit information on their expertise and credentials via this online form.
- CCF Experts approve 3-5 experts per organ and give them access to the ASCT+B table forms so they can author and review the forms.

Construction by Subject Matter Experts (SMEs)

- MC-IU provide pre-populated initial ASCT+B table with UBERON and CL ontology IDs.
- A first set of organ experts authors the tables and indicates author contributions.
- Authors use the ASCT+B Reporter to identify/resolve naming and interlinkage issues.
- Completed tables are submitted to the CCF Experts for review.

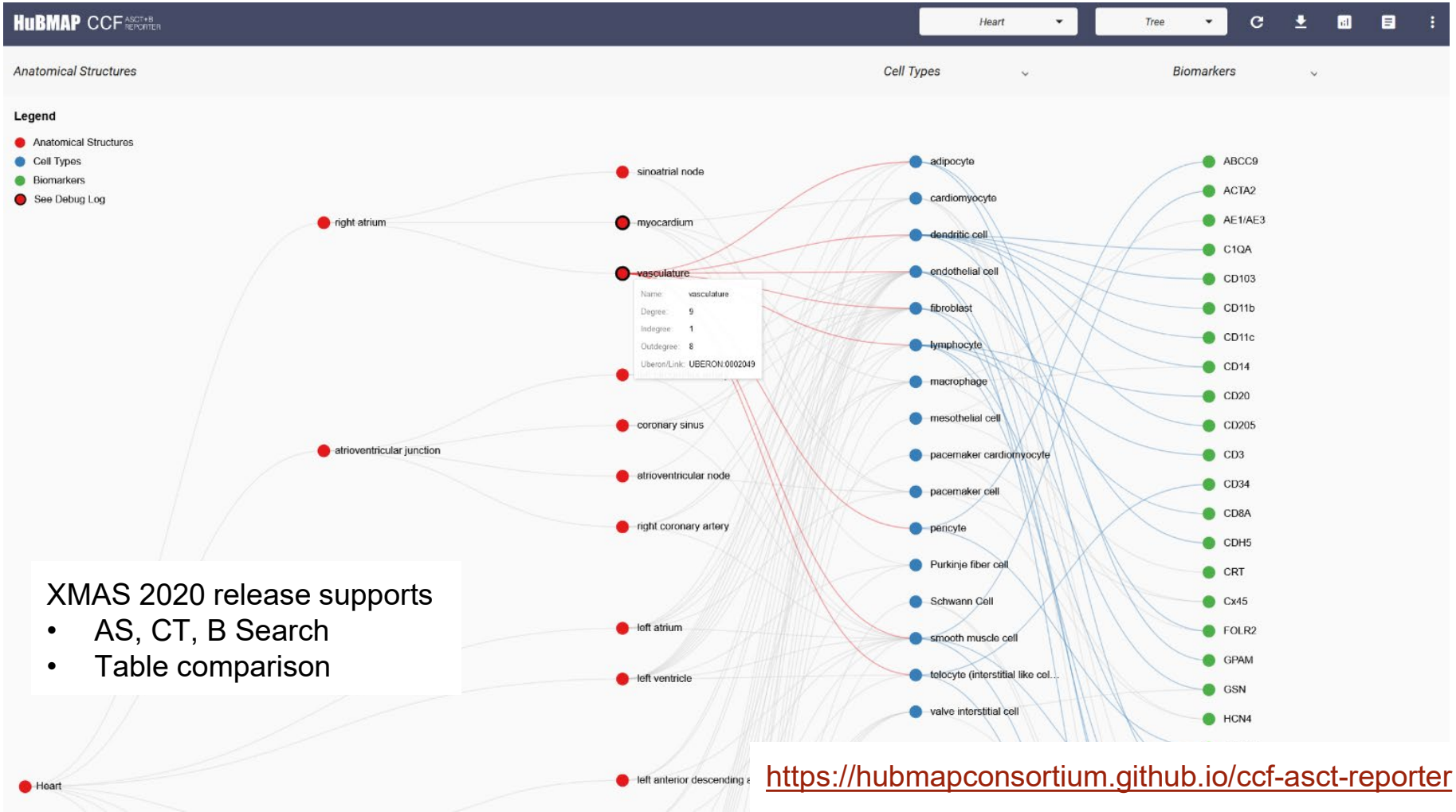
Review by Subject Matter Experts (SMEs)

- The beginning of each month, all tables ready for review are submitted by CCF Experts to a second set of organ experts for review.
- Review criteria include: scientific rigor (citation of publications, data), coverage and quality of the ASCT+B tables.
- Review results comprise detailed comments together with a rating (accepted, accepted with minor or major revisions, rejected) and are shared back with the author team.

Review by CCF Experts

The begin of each month, all tables ready for review are cross-checked against

1. existing ontologies, e.g., UBERON, CL, to identify any terms that might be missing or that might have different spelling. The goal is to arrive at ASCT+B tables that are in close alignment with existing ontologies so only few changes need to be requested from ontology owners.



XMAS 2020 release supports

- AS, CT, B Search
- Table comparison

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

Table S1. Canonical cell types (45) in the human lung and their abundances, markers, and available expression data.

Cell type	Relative abundance (%)	Number (millions)[a]	Canonical markers[b]	Extant expression profiles: Single cell	Extant expression profiles: Primary[c]	Expression accession codes	Abundance reference (method)[d]
Epithelium							
Club Cell	0.5	1,500	CYP2F2, SCGB3A2, CCKAR	Yes	Yes	MTAB-6149, E-MTAB-6653	Boers et al. 1999 (e)
Ciliated Cell	2	6,000	FOXJ1, TUBB1, TP73, CCDC78	Yes	Yes	GSE122960	Raman et al. 2009 (e)
Basal Cell	0.5	1,500	KRT5, KRT14, TP63, DAPL1	Yes	Yes	MTAB-6149, E-MTAB-6653	Boers et al. 1998 (e)
Goblet Cell	0.2	500	MUC5B, MUC5AC, SPDEF	Yes	Yes	EGAS00001001755	Boers et al. 1999 (e)
Mucous Cell	0.03	80	MUC5B				Widdicombe and Wine 2015 (e)
Serous Cell	0.03	80	PRR4, LPO, LTF	Yes	Yes	EGAS00001001755	Basbaum et al. 1990 (e)
Ionocyte	0.03	100	CFTR, FOXI1, ASCL3	Yes	Yes	EGAS00001001755	Montoro et al. 2018 (e)
Neuroendocrine Cell	0.01	40	CALCA, CHGA, ASCL1	Yes	Yes	EGAS00001001755	Boers et al. 1996 (e)
Tuft Cell	0.1	200	DCLK1, ASCL2	Yes		GSE102580	Chang et al. 1986; Montoro et al. 2018 (e)
Alveolar Epithelial Type 1 Cell	13	40,000	AGER, PDPN, CLIC5	Yes	Yes	MTAB-6149, E-MTAB-6653	Crapo et al. 1982 (f)
Alveolar Epithelial Type 2 Cell	7	20,000	SFTPB, SFTPC, SFTPD, MUC1, ETV5	Yes	Yes	GSE122960	Crapo et al. 1982; Fehrenbach et al. 1994 (f)
Total	23	70,000					
Endothelium							
Artery Cell	1	3,000	GJA5, BMX	(bulk)	(cultured)	phs000998.v1.p1	Townsley et al. 2012; The Lung, Chapter 74 (g)
Vein Cell	1	3,000	ACKR1				Townsley et al. 2012; The Lung, Chapter 74 (g)
		70,000	CA4				Crapo et al. 1982 (f)
		2,000					Deffebach et al. 1987 (g)
		2,000	PROX1, PDPN	Yes	Yes	MTAB-6149, E-MTAB-6653	Kambouchner et al. 2009; Sozio et al. 2012 (g)
		80,000		(unannotated)		MTAB-6149, E-MTAB-6653, EGAS00001001755	
		5,000	CNN1, ACTA2, TAGLN, RGS5	Yes	Yes	GSE75990	Townsley et al. 2012; The Lung, Chapter 74 (h)
		4,000	CNN1, ACTA2, TAGLN, DES, LGR6			GSE75990	Elliot et al. 1999; The Lung, Chapter 74 (h)
		20,000	COL1A1, PDGFRA	Yes	Yes	EGAS00001001755	Crapo et al. 1982 (f,i)
		20,000	COL1A1, PDGFRA, ELN, ACTA2	Yes	Yes	EGAS00001001755	Crapo et al. 1982 (f,i)
		20,000	COL1A1, PDGFRA, PLIN2, APOE				Crapo et al. 1982 (f,i)
Pericyte	7	20,000	CSPG4, TRPC6, PDGFRB	(bulk)	(cultured)	GSE75990	Crapo et al. 1982 (f,i)
Mesothelial Cell	0.3	1,000	MSLN, UPK3B, WT1	(bulk)	(cultured)	GSE63966	Michailova et al. 1997 (j)
Total	30	90,000		(unannotated)		MTAB-6149, E-MTAB-6653, EGAS00001001755	
PNS							
Intrinsic Neuron	0.0003	1	SNAP25				Fox et al. 1980; Sparrow et al. 1999 (j)
Glial Cell	0.0002	0.5					Sparrow et al. 1999 (j)
Total	0.0005	1.5					
Immune							
B Cell	0.5	1,500	CD79A, CD24, MS4A1, CD19	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Finkelstein et al. 1995; Banat et al. 2015 (k)
Plasma Cell	0.7	2,000	CD79A, CD27, SLAMF7	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Banat et al. 2015 (k)
CD8+ Mem/Eff T Cell	1	3,000	CD3E, CD8A, GZMK, DUSP2	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Finkelstein et al. 1995; Banat et al. 2015 (k,i)
CD8+ Naive T Cell	1	3,000	CD3E, CD8, GZMH, GZMB	Yes	Yes	MTAB-6149, E-MTAB-6653	Finkelstein et al. 1995; Banat et al. 2015 (k,l)
CD4+ Mem/Eff Cell	0.7	2,000	CD3E, CD8, COTL1, LDHB	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Finkelstein et al. 1995; Banat et al. 2015 (k,i)
CD4+ Naive T Cell	0.7	2,000	CD3E, CD4, CCR7, LEF1	Yes	Yes	MTAB-6149, E-MTAB-6653	Finkelstein et al. 1995; Banat et al. 2015 (k,i)
Natural Killer Cell	1	3,000	KLRD1, NKG7, TYROBP	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Marquardt et al. 2017 (l)
Natural Killer T Cell	0.7	2,000	CD3E, CD8A, FCER1G, TYROBP	Yes	Yes	MTAB-6149, E-MTAB-6653	Marquardt et al. 2017 (k,i)
Neutrophil	0.8	2,500	S100A8, S100A9, IFITM2, FCGR3B	Yes	Yes	EGAS00001001755	Finkelstein et al. 1995; Banat et al. 2015 (k)
Basophil	0.3	1,000	MS4A2, CPA3, TPSAB1	Yes	Yes	MTAB-6149, E-MTAB-6653	Finkelstein et al. 1995 (k,i)
Mast Cell	1	3,000	MS4A2, CPA3, TPSAB1	Yes	Yes	MTAB-6149, E-MTAB-6653	Finkelstein et al. 1995; Banat et al. 2015 (k)
Eosinophil	0.3	1,000	SIGLEC8	(bulk)	(cultured)		Finkelstein et al. 1995 (k,i)
Megakaryocyte	0.3	1,000	NRGN, PPBP, PF4, OST4	(bulk)	Yes		Dejima et al. 2018; Skoczynski et al. 2019 (m)
Macrophage	7	20,000	MARCO, MSR1, MRC1	Yes	Yes	MTAB-6149, E-MTAB-6653	Crapo et al. 1982; Fehrenbach et al. 1994 (f)
Plasmacytoid Dendritic Cell	0.3	800	LILRB4, IRF8, LILRA4	Yes	Yes	GSE94820	Banat et al. 2015 (k,i)
Myeloid Dendritic Cell 1	0.3	1,000	MHCII CLEC9A, LAMP3	Yes	Yes	GSE94820	Banat et al. 2015 (k)
Myeloid Dendritic Cell 2	0.1	200	MHCII, CD1C, PLD4	Yes	Yes	GSE94820	Banat et al. 2015 (k)
Classical Monocyte	2	4,000	CD14, S100A8	Yes	Yes	E-MTAB-6701, E-MTAB-6678	Hance et al. 1985; Hoogsteden et al. 1989 (k,i)
Intermediate Monocyte	2	4,000	CD14, S100A8, CD16	(bulk)	Yes	GSE80095	Hance et al. 1985; Hoogsteden et al. 1989 (k,i)
Nonclassical Monocyte	1	3,000	CD16	Yes	Yes	GSE94820	Hance et al. 1985; Hoogsteden et al. 1989 (k)
Total	20	60,000					
Total (all compartments)	100	300,000					

a, numbers of each type were calculated with their abundances and the total number of lung cells (estimated by comparing volume of lungs to the whole body). **b,** Canonical markers were obtained from referenced expression data or commonly used markers in the literature. **c,** Expression profiles captured immediately following tissue dissociation are considered primary. **d,** Alveoli were assumed to occupy ~90% of the total lung volume for all estimations. **e,** Inferred from mean relative abundance in proximal, medial and distal airway epithelium. **f,** Calculated by stereology **g,** Resin casts showed similar surface area of arteries and veins. **h,** Vascular smooth muscle is estimated to be slightly more abundant than airway smooth muscle. **i,** abundance of a more general cell type was split evenly. **j,** Inferred from impression of light or electron microscopy. **k,** Inferred from histological abundance in non-perfused healthy tissue. **l,** Inferred from abundance among immune cells with FACS. **m,** Calculated using microfluidic capture

CTs

Bs

Refs

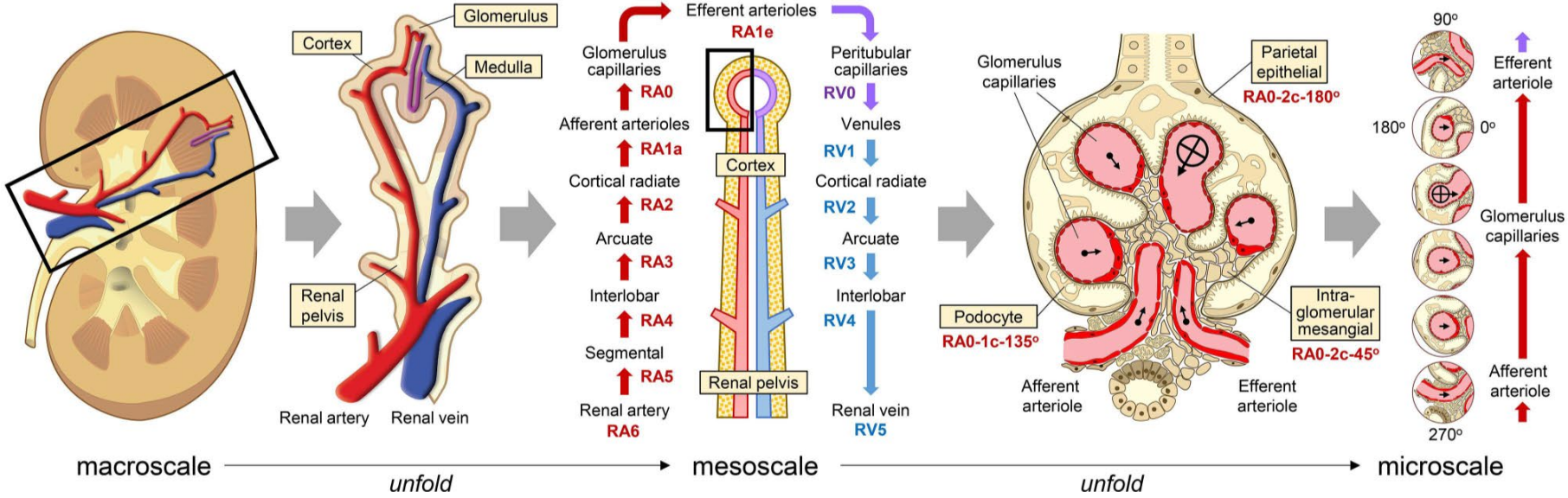
A molecular cell atlas of the human lung from single cell RNA sequencing

Kyle J. Travaglini, Ahmad N. Nabhan, Lolita Penland, Rahul Sinha, Astrid Gillich, Rene V. Sit, Stephen Chang, Stephanie D. Conley, Yasuo Mori, Jun Seita, Gerald J. Berry, Joseph B. Shrager, Ross J. Metzger, Christin S. Kuo, Norma Neff, Irving L. Weissman, Stephen R. Quake, Mark A. Krasnow

doi: https://doi.org/10.1101/742320

As are in Table S2

Capturing vasculature details from macro to micro scale is critically important for a vasculature based CCF



Weber, Griffin M, Yingnan Ju, and Katy Börner. 2020. "Considerations for Using the Vasculature as a Coordinate System to Map All the Cells in the Human Body". *Frontiers in Cardiovascular Medicine* 7 (29): doi: 10.3389/fcvm.2020.00029.

ASCT+B Table Example: Kidney vasculature

Vasculature								
	renal atery [L/R]					Endothelial Cell (general)	EC	*EMCN*, PECAM1*, FLT1**
		segmental arteries [superior, inferior, anterior, posterior]						
		interlobar aterties						
		arcuate aterties						
		cortical radiate ateries {cortex}	afferent arterioles {nephron}			EC-Afferent/Efferent Arteriole	EC-AEA	*SERPINE2*, TM4SF1**
				glomerulus capillaries {glomerulus}		Capillary Endothelial Cell	GC-EC	*EHD3*, EMCN*, HECW2*, FLT1*, AQP1**
			efferent arterioles {nephron}			EC-Afferent/Efferent Arteriole	EC-AEA	*SERPINE2*, TM4SF1**
				peritubular capillaries		EC-Peritubular capillaries	EC-PTC	*PLVAP**
				descending vasa recta		EC-Descending Vasa Recta	EC-DVR	*TM4SF1*, PALMD*
				ascending vasa recta		EC-Ascending Vasa Recta	EC-AVR	*DNASEIL3**
	renal vein [L/R]	cortical radiate veins {cortex}	venules			Endothelial Cell (general)	EC	*EMCN*, PECAM1*, FLT1**
		arcuate veins						
		interlobar veins						

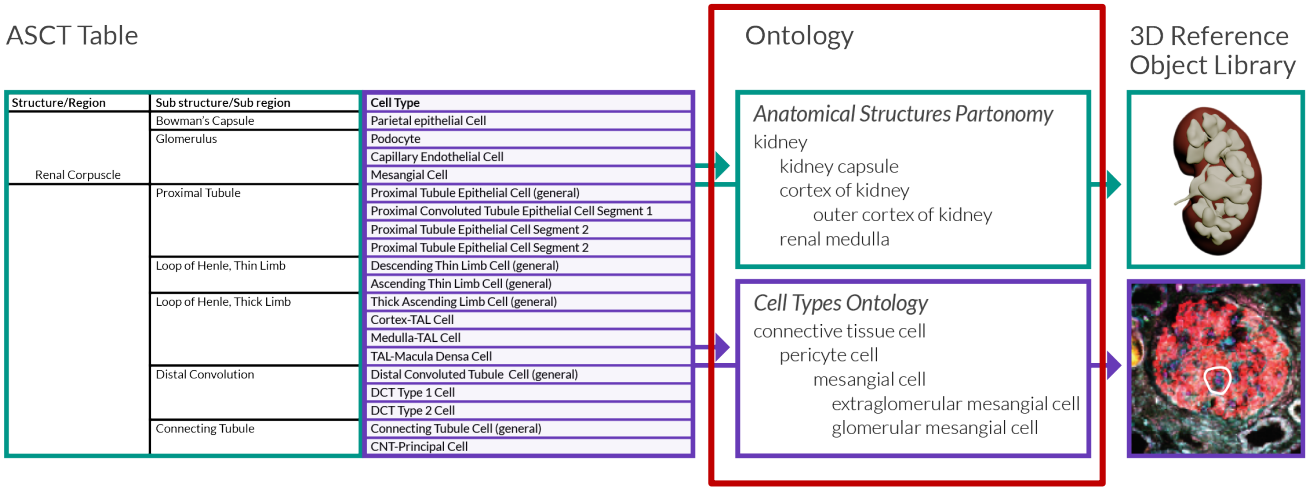
Vasculature	renal artery [L/R]				Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal artery [L/R]	segmental arteries [superior, inferior, anterior, posterior]			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal artery [L/R]	interlobar arterties			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal artery [L/R]	arcuate arteries			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	afferent arterioles {nephron}		EC-Afferent/Efferent Arteriole	EC-AEA	SERPINE2*, TM4SF1*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	afferent arterioles {nephron}	glomerulus capillaries {glomerulus}	Capillary Endothelial Cell	GC-EC	EHD3*, EMCN*, HECW2*, FLT1*, AQP1*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	efferent arterioles {nephron}		EC-Afferent/Efferent Arteriole	EC-AEA	SERPINE2*, TM4SF1*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	efferent arterioles {nephron}	peritubular capillaries	EC-Peritubular capillaries	EC-PTC	PLVAP*
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	efferent arterioles {nephron}	descending vasa recta	EC-Descending Vasa Recta	EC-DVR	TM4SF1*, PALMD
Vasculature	renal artery [L/R]	cortical radiate arteries {cortex}	efferent arterioles {nephron}	ascending vasa recta	EC-Ascending Vasa Recta	EC-AVR	DNASEIL3*
Vasculature	renal vein [L/R]	cortical radiate veins {cortex}	venules {nephron}		Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal vein [L/R]	cortical radiate veins {cortex}			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal vein [L/R]	arcuate veins			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*
Vasculature	renal vein [L/R]	interlobar veins			Endothelial Cell (general)	EC	EMCN*, PECAM1*, FLT1*

Weber, Griffin M, Yingnan Ju, and Katy Börner. 2020. "Considerations for Using the Vasculature as a Coordinate System to Map All the Cells in the Human Body". *Frontiers in Cardiovascular Medicine* 7 (29): doi: 10.3389/fcvm.2020.00029.

ASCT+B Table Usage

ASCT+B Table Usage

ASCT+B tables guide **CCF Ontology** and **3D Reference Object Library** design that semantically name and spatially place tissue data from different donors into one CCF (i.e., mapping).

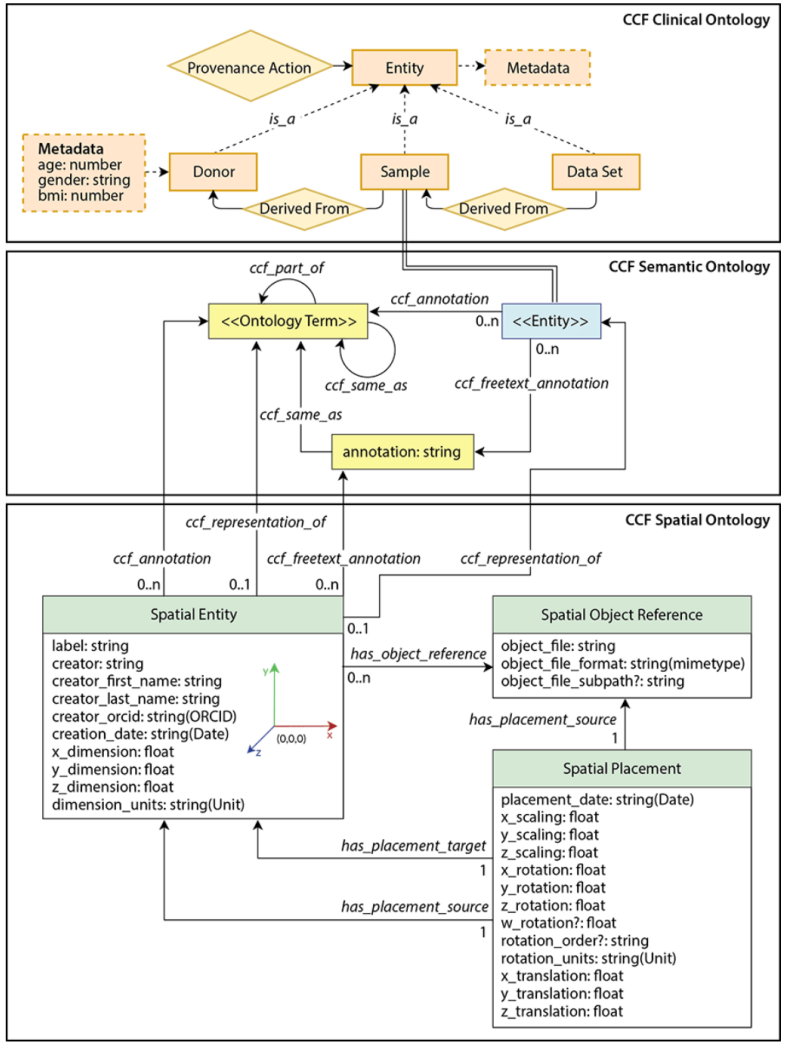


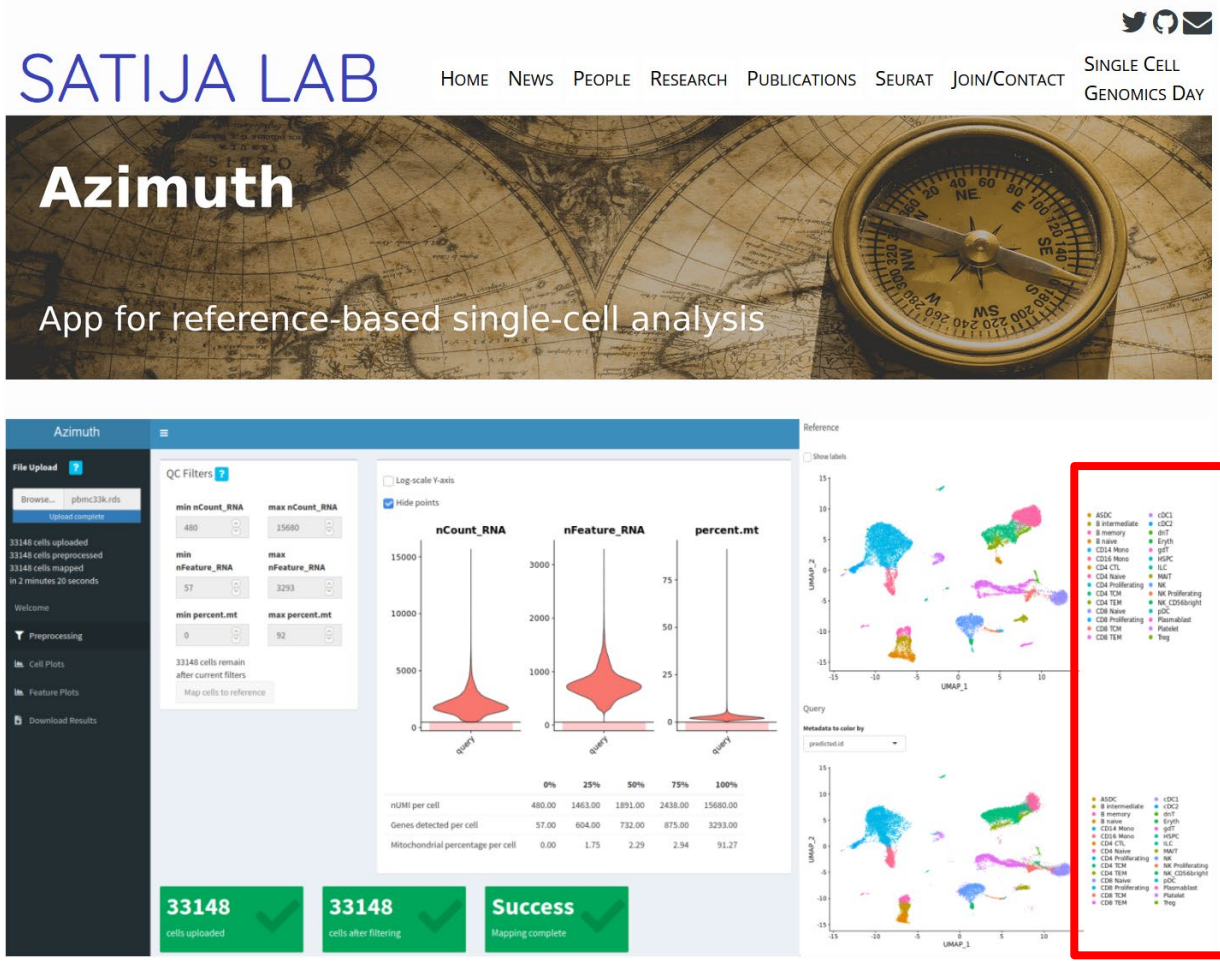
Tissue blocks are registered into the CCF using the Registration User Interface (RUI), and they can be explored via the Exploration User Interface (EUI).

CCF Ontology v1.5.0

References

- Herr II, BW and Börner K. HuBMAP Common Coordinate Framework. https://bioportal.bioontology.org/ontologies/CCF/
- Herr II, BW, Quardokus EM, Cross LE, Record EG, Weber GM, and Börner K. HuBMAP CCF Ontology Source Code Repository.
- Börner K, Quardokus EM, Herr II, BW, Cross LE, Record EG, Ju Y, Bueckle A, Sluka JP, Silverstein J, Browne K, Jain S, Wasserfall CH, Jorgensen ML, Spraggins JM, Patterson NH, Weber GM. 2020. Construction and Usage of a Human Body Common Coordinate Framework Comprising Clinical, Semantic, and Spatial Ontologies. https://arxiv.org/abs/2007.14474.



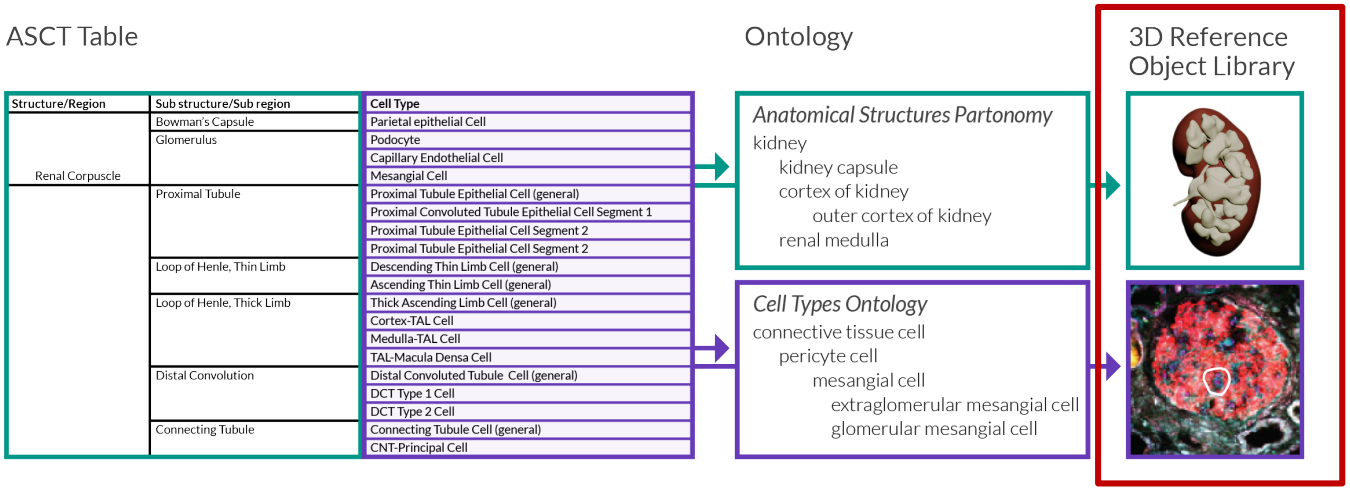


https://satijalab.org/azimuth

CT terms from ASCT+B linked to Cell Ontology

ASCT+B Table Usage

ASCT+B tables guide **CCF Ontology** and **3D Reference Object Library** design that semantically name and spatially place tissue data from different donors into one CCF (i.e., mapping).



Tissue blocks are registered into the CCF using the Registration User Interface (RUI), and they can be explored via the Exploration User Interface (EUI).

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

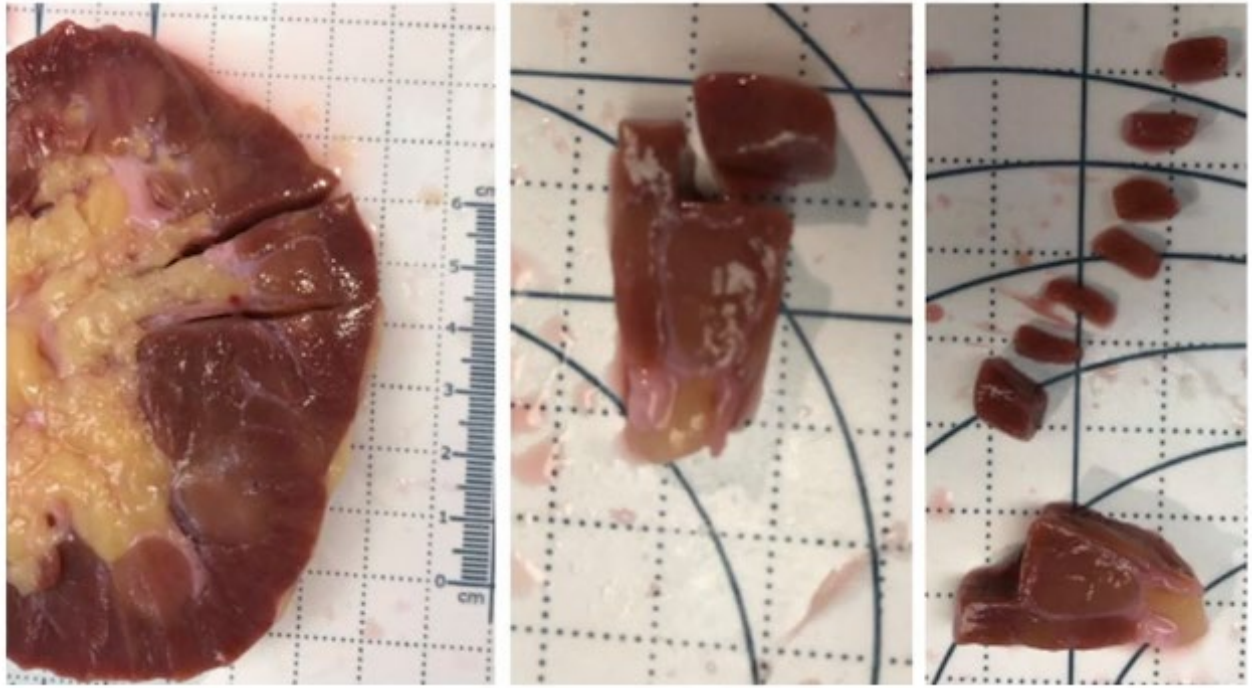
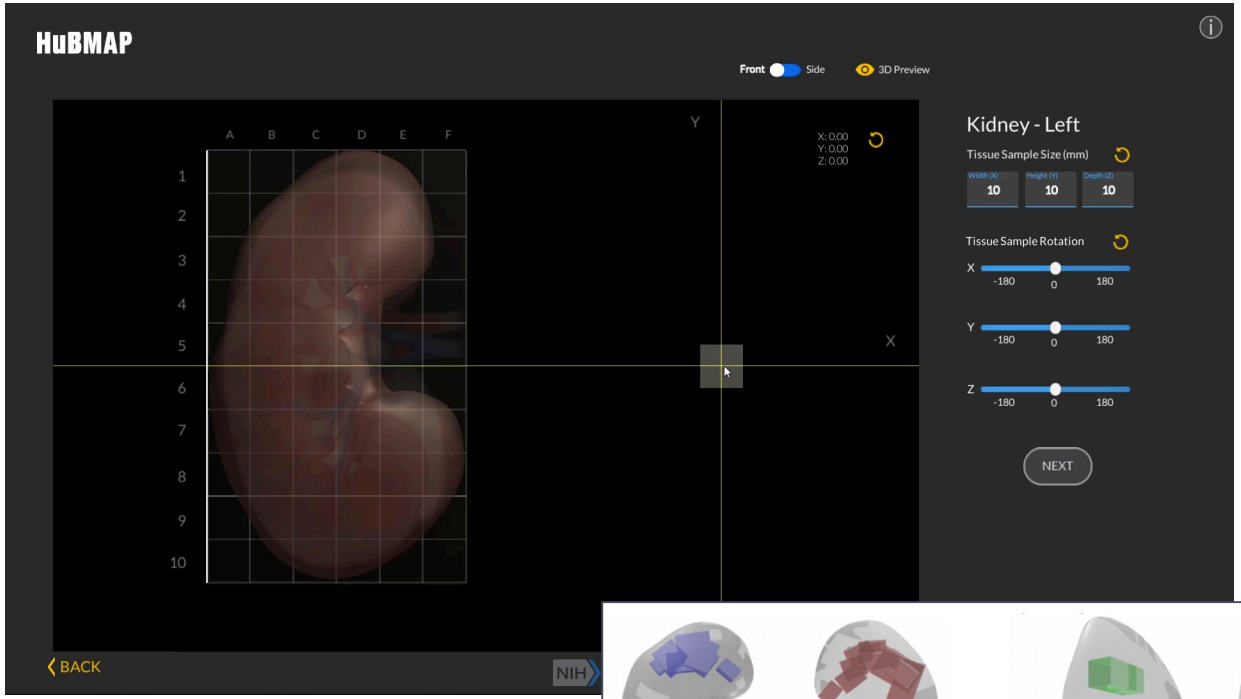
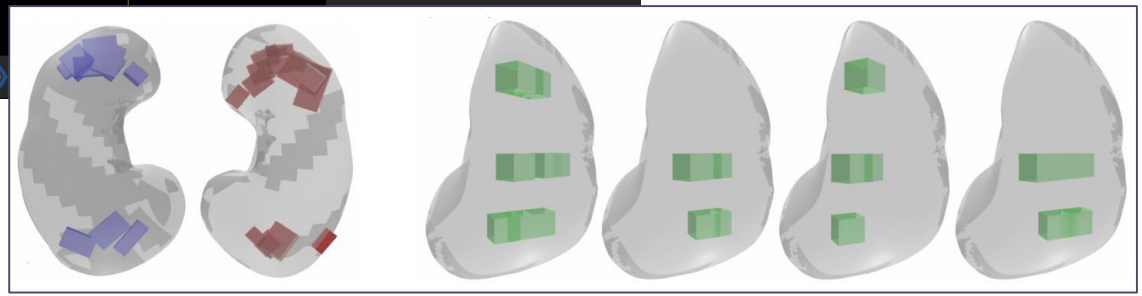
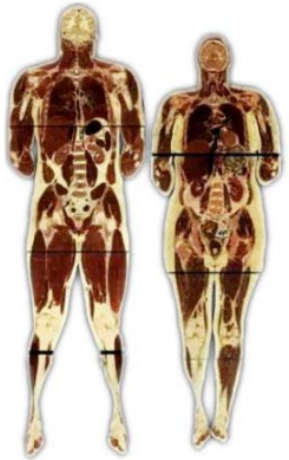


Image provided by Sanjay Jain, TMC-UCSD

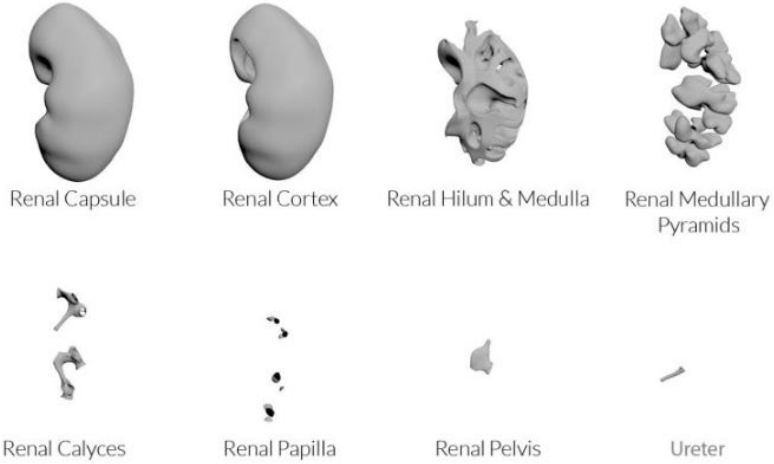


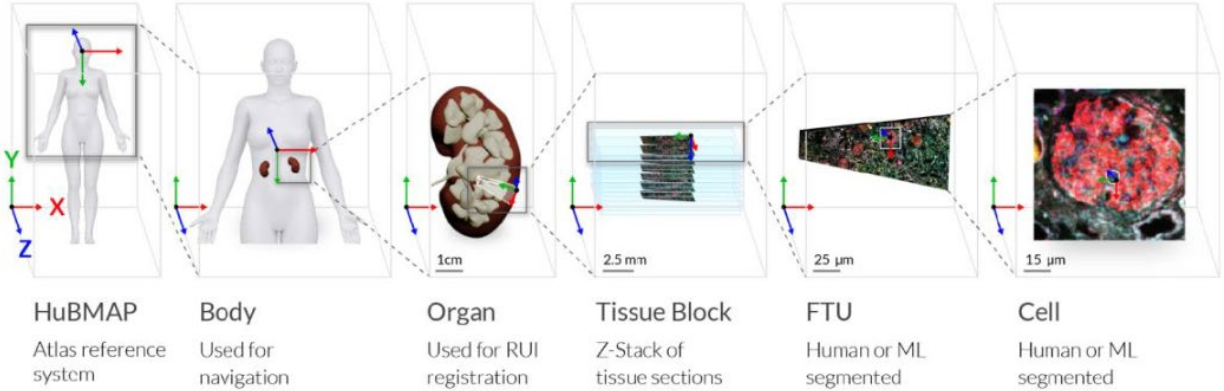


For the first HuBMAP portal release, 48 tissue blocks were registered.



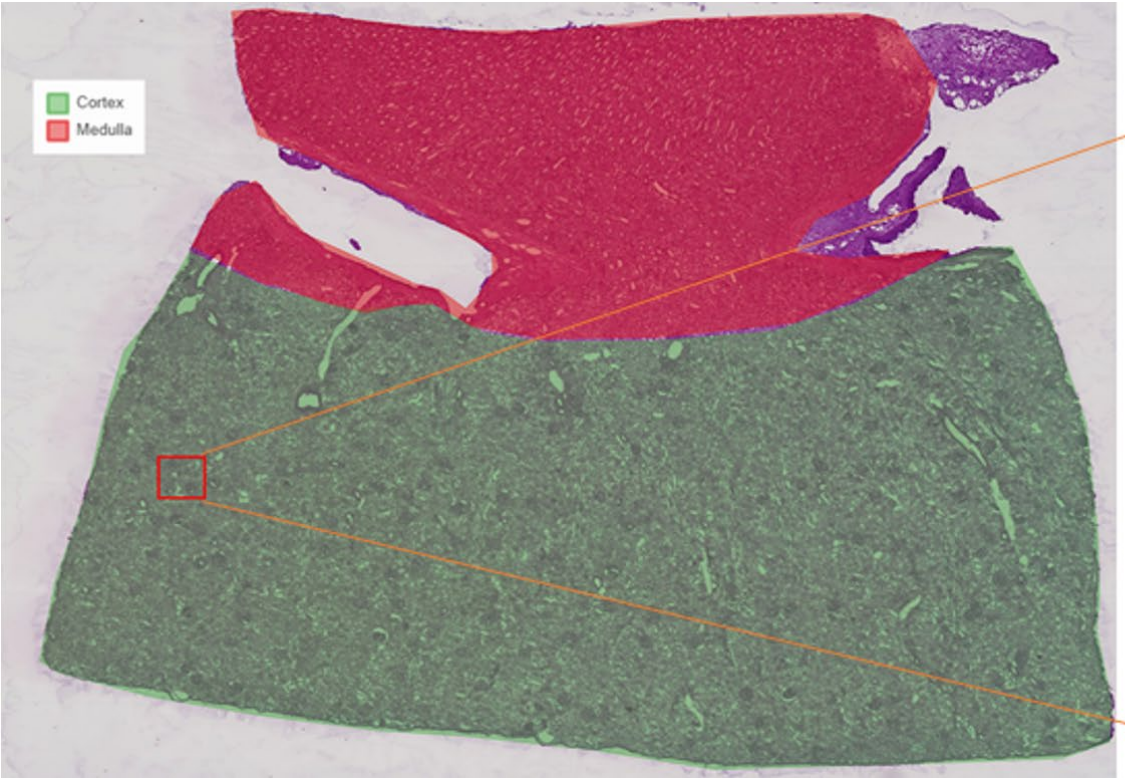
NLM Visible Human

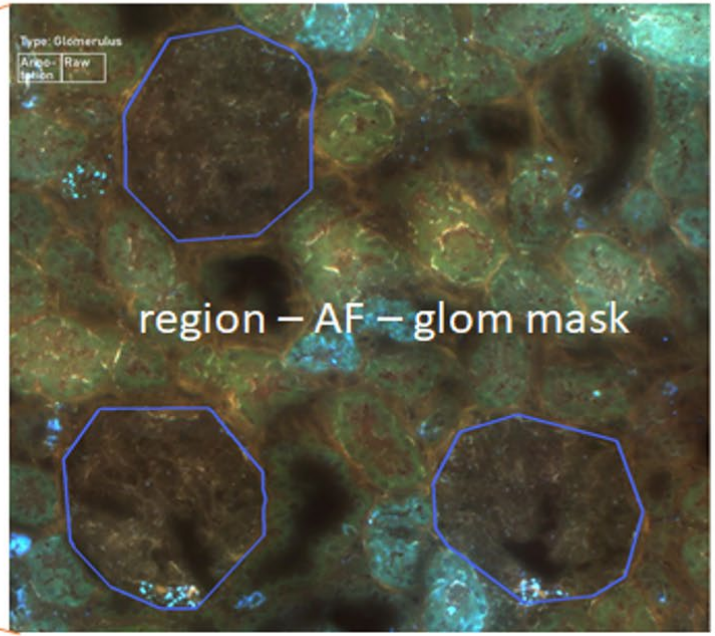




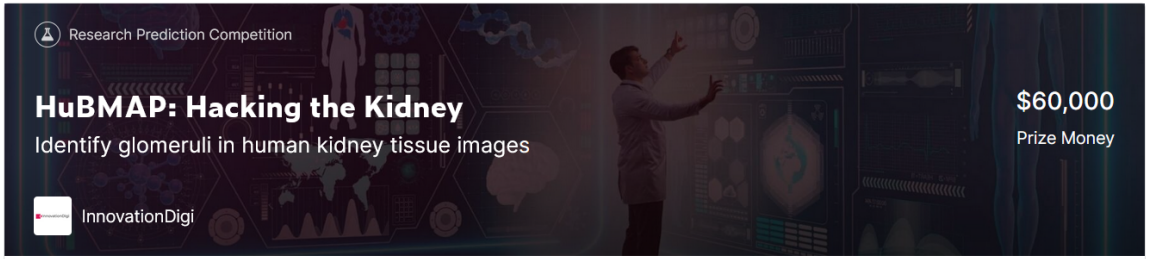
https://hubmapconsortium.github.io/ccf/pages/ccf-3d-reference-library.html

Public private partnership with NIH, Google, Lilly and other sponsors.





https://www.kaggle.com/c/hubmap-kidney-segmentation



Overview Data Notebooks Leaderboard Rules Team Host My Submissions

This competition is not yet live; only competition hosts can currently view it.

Overview Edit

- Description
- Supervised ML Evaluation
- Judges Prize
- Prizes
- Timeline
- Organizers & Sponsors

+ Add Page

Our best estimates show there are over 7 billion people on the planet and 300 billion stars in the Milky Way. By comparison, the adult human body contains 37 *trillion* cells. To determine the function and relationship among these cells is a monumental undertaking. Many areas of human health would be impacted if we better understand cellular activity. A problem with this much data is a great match for the Kaggle community.

Just as the Human Genome Project mapped the entirety of human DNA, the Human BioMolecular Atlas Program (HuBMAP) is a major endeavor. Sponsored by the National Institutes of Health (NIH), HuBMAP is working to catalyze the development of a framework for mapping the human body at a level of glomeruli functional tissue units for the first time in history. Hoping to become one of the world's largest collaborative biological projects, HuBMAP aims to be an open map of the human body at the cellular level.

This competition, "Hacking the Kidney," starts by mapping the human kidney at single cell resolution.

Your challenge is to detect functional tissue units (FTUs) across different tissue preparation pipelines. An FTU is defined as a "three-dimensional block of cells centered around a capillary, such that each cell in this block is within diffusion distance from any other cell in the same block" (de Bono, 2013). The goal of this competition is the implementation of a successful and robust glomeruli FTU detector.

You will also have the opportunity to present your findings to a panel of judges for additional consideration. Successful submissions will construct the tools, resources, and cell atlases needed to determine how the relationships between cells can affect the health of an individual.

https://www.kaggle.com/c/hubmap-kidney-segmentation

OUR JUDGES

THOMAS FUCHS
Founder and CSO of Paige. or AI, Director at **Memorial Sloan Kettering Cancer Center**, Professor at **Weill Cornell**

AMY BERNARD
Director, Science & Technology Strategy, **Allen Institute**

MAIGAN BRUSKO
Department of Pathology, Immunology, and Laboratory Medicine at the **University of Florida**

JOHN MARIONI
Research Group Leader, **EMBL-EBI**

ZORINA GALIS
Chief, Vascular Biology and Hypertension, National Heart Lung and **Blood Institute/National Institutes of Health, Bethesda, MD**

BLUE LAKE
Integrative Genomics Group
Department of Bioengineering
University of California at San Diego

DAVID VAN VALEN
Division of Biology and Bioengineering, **California Institute of Technology**

MATT NELSON
Vice Precedent, Genetics and Genomics at **Deerfield**

ANDY PALMER
CEO and co-founder of **Tamr Inc.**

LUCY COLWELL
Research Scientist at **Google**

ALEX WOLF
Head of Applied Machine Learning at **Cellarity**

https://www.kaggle.com/c/hubmap-kidney-segmentation

SOP for Approval of 3D Reference Objects

CCF 3D Reference Object Library

Overview

The CCF 3D Reference Object Library provides anatomically correct reference organs. The organs are developed by a specialist in 3D medical illustration and approved by organ experts, see SOP.

Initially, reference objects were created using data from the Visible Human male and female datasets provided by the National Library of Medicine. The male dataset comprises 1,871 cross-sections at 1mm intervals for both CT and anatomical images at a resolution of 4,096 pixels by 2,700 pixels. The female data set has the same characteristics as the Visible Human Male but axial anatomical images were obtained at 0.33 mm intervals resulting in 5,189 cross-section anatomical images. The male was white, 180.3 cm (71 inch) tall, 199-pound and was 38 years old. The female was white, 171.2 cm (67.4 inch) tall, obese, and 59 years old.

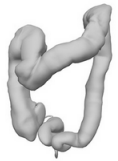
For the 1st HuBMAP Portal Release, kidney and spleen reference organs are freely available in GLB format. They can be viewed and explored using free web browsers such as Babylon.js. Screenshots and major properties of the nested reference organ objects are given in table below.

For selected organs, 3D extraction site objects are provided. Some extraction sites resemble geometric objects (e.g., cuboids for heart) while others take the shape of one or more whole or partial anatomical structures (e.g., in spleen). The 3D extraction sites do not restrict registration to specific regions, instead they provide "expert defined landmarks" to help guide tissue registration. The extraction site objects are also used for automatic semantic annotation of tissue samples via collision detection during registration.

Reference Organs

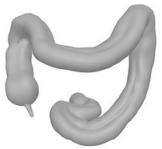
COLON | HEART | KIDNEY | SPLEEN

MALE: Colon



# Anatomical Structures	10
Appendix	1
Ascending Colon	1
Cecum	1
Descending Colon	1
Hepatic Flexure	1
Ileocecal Valve	1
Rectum	1
Sigmoid Colon	1
Splenic Flexure	1
Transverse Colon	1

FEMALE: Colon



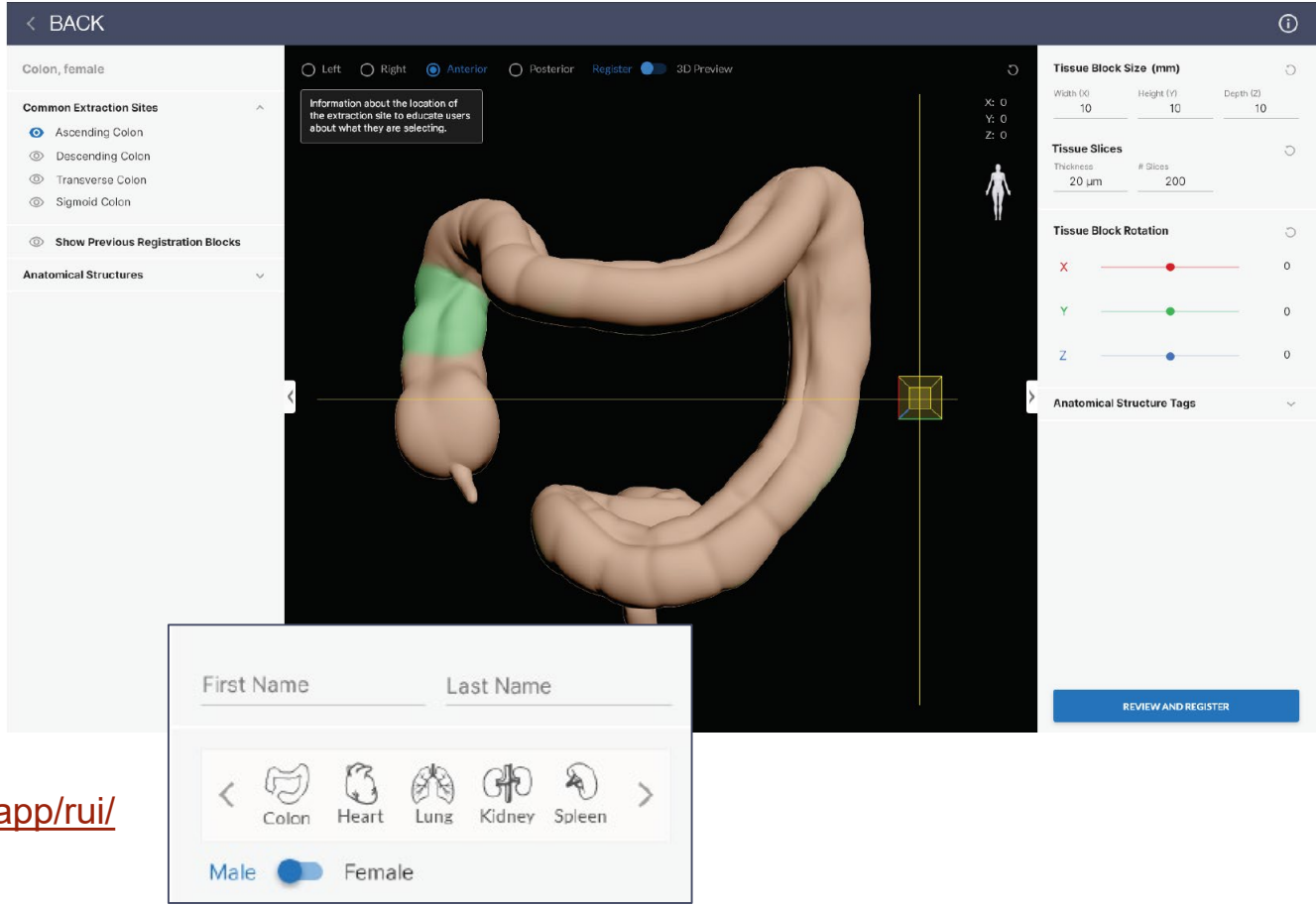
# Anatomical Structures	10
Appendix	1
Ascending Colon	1
Cecum	1
Descending Colon	1
Hepatic Flexure	1
Ileocecal Valve	1
Retum	1
Sigmoid Colon	1
Splenic Flexure	1
Transverse Colon	1

https://hubmapconsortium.github.io/ccf/dld/SOP-3D-Reference-Object-Approval-v1.0.1.pdf

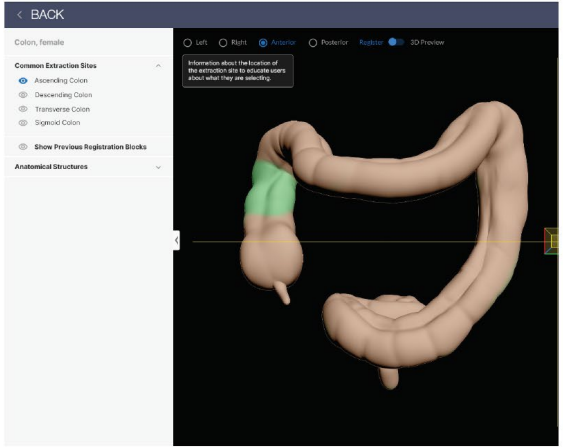
CCF Registration User Interface (RUI) v1.0.0

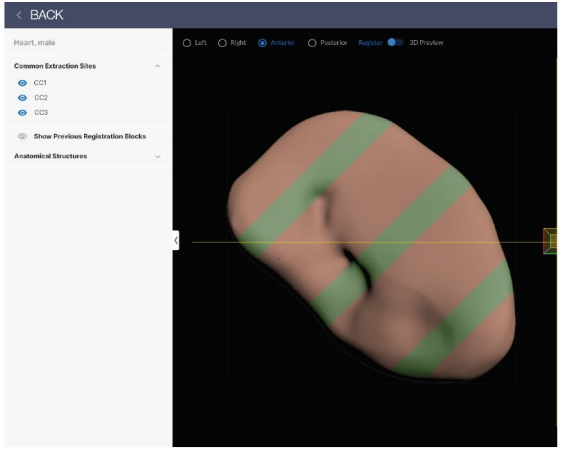
New Features:

- 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

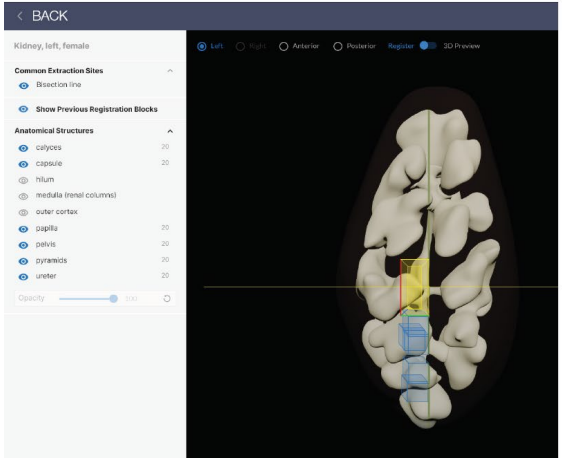
Spleen

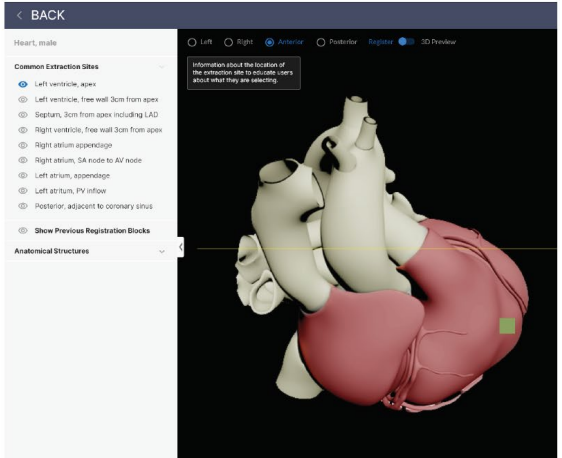
- CC1
- CC2
- CC3

Colon

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

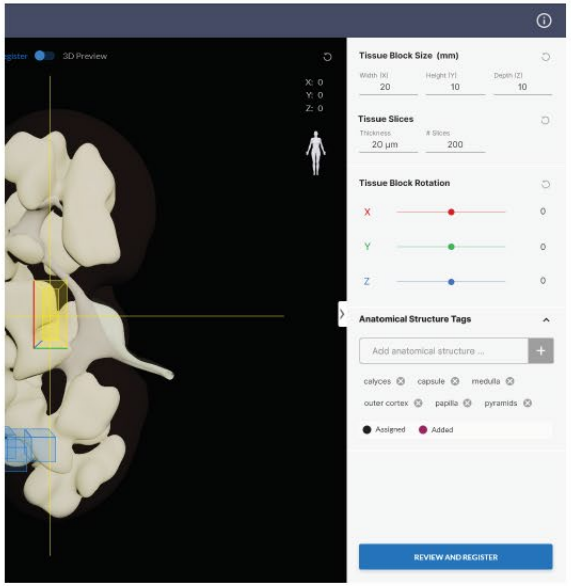
Heart	Extraction Site Mapping
Left atrium, appendage	7
Left atrium, PV inflow	8
Left ventricle, apex	1
Left ventricle, free wall 3cm from apex	2
Septum, 3cm from apex including LAD	3
Posterior, adjacent to coronary sinus	9
Right atrium appendage	5
Right atrium, AV (atrioventricular) node	6a
Right atrium,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



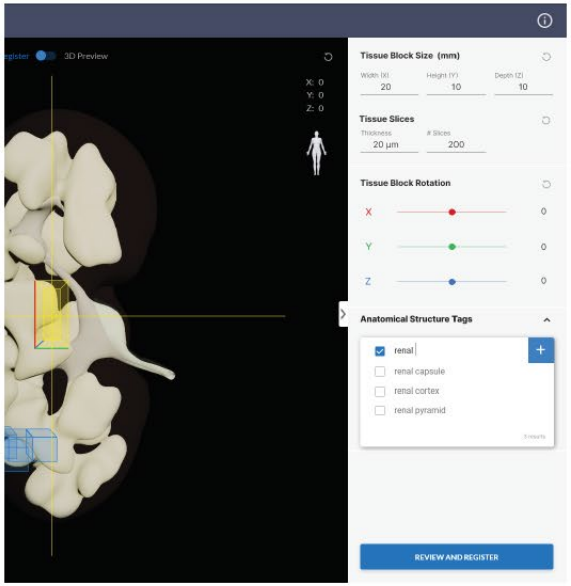


CCF Registration User Interface (RUI) v1.0.0 cont.

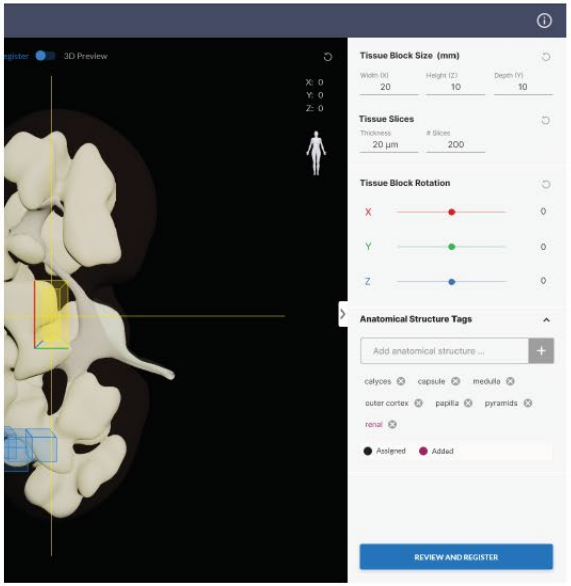
Collision when Tissue Block hits Reference Organ



Tag Search behavior



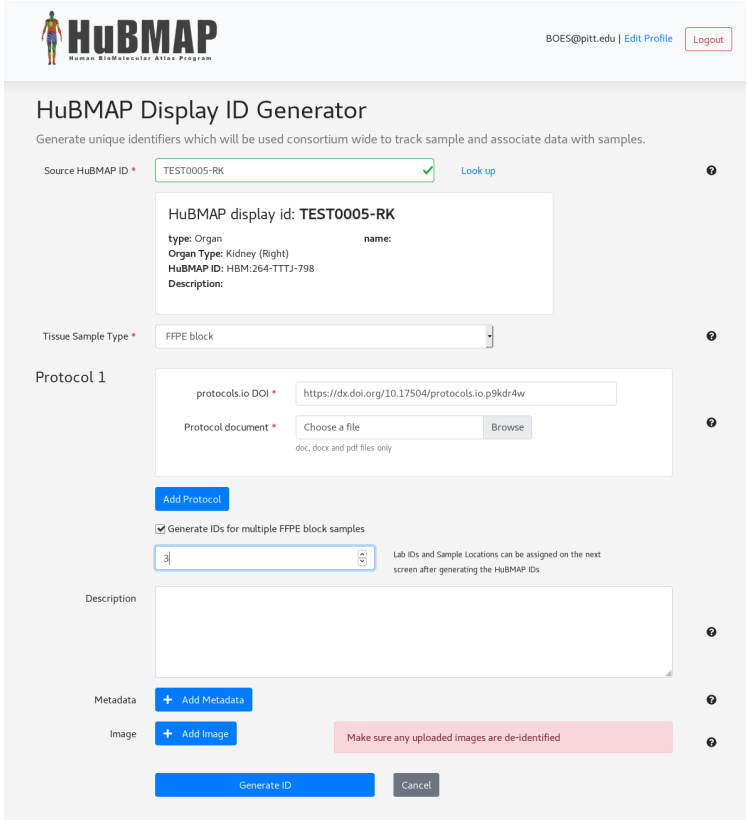
Custom tag added to list

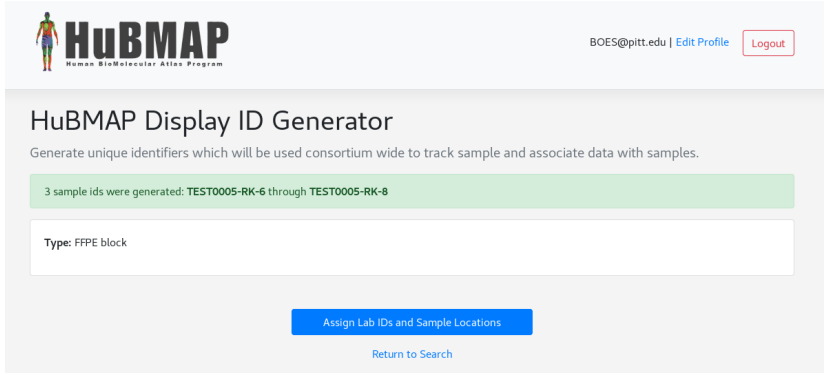


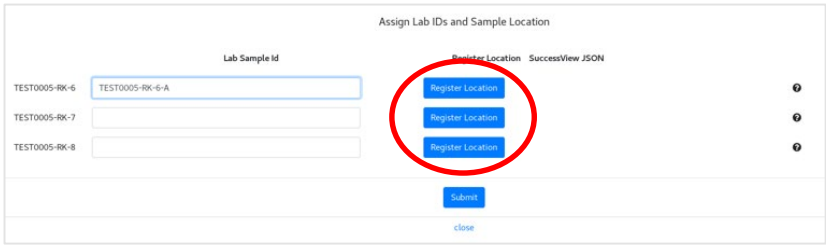
https://hubmap-ccf-ui.netlify.app/rui/

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

HuBMAP Upload Portal

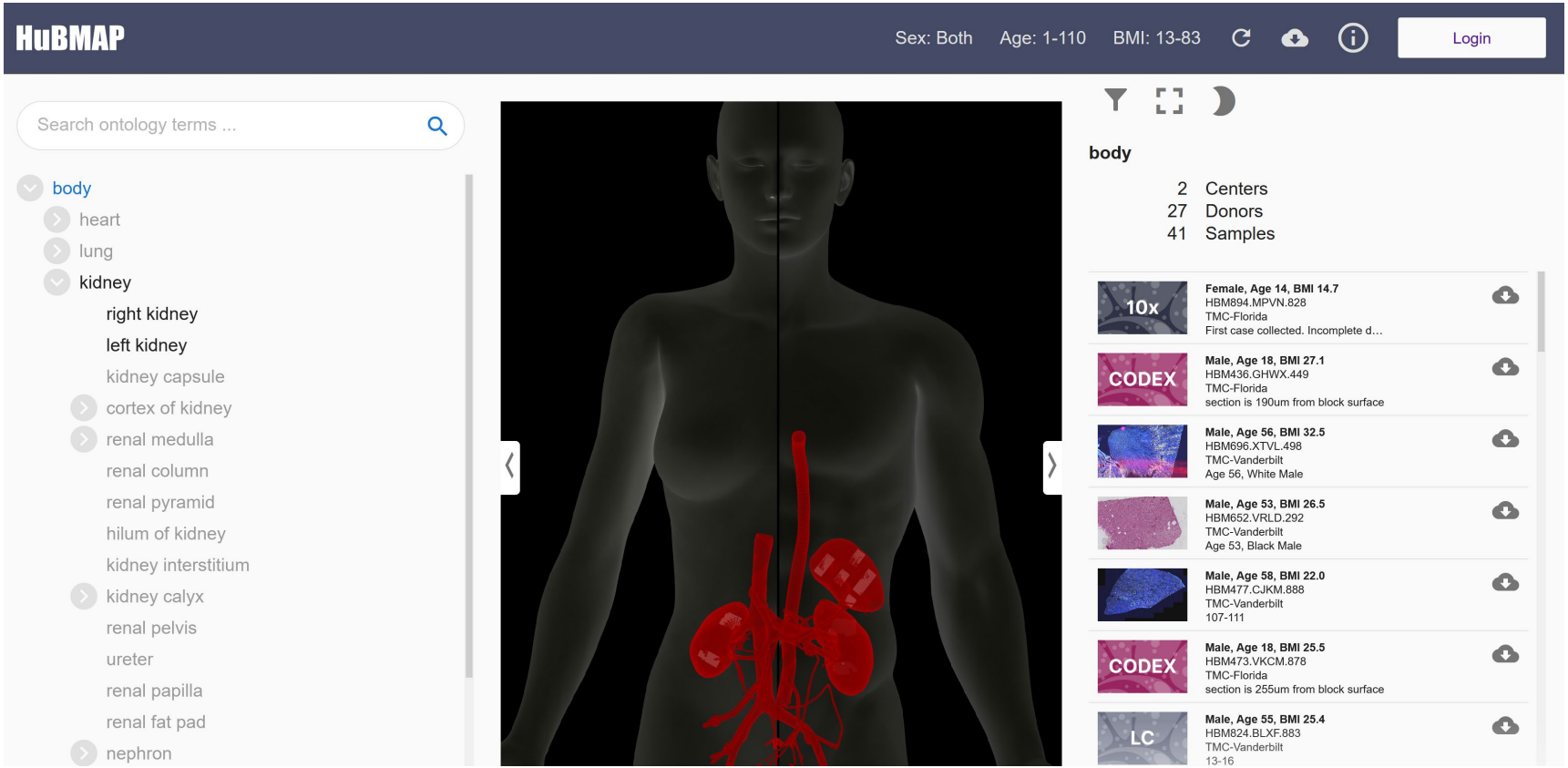






Implemented by the HIVE IEC

CCF Exploration User Interface (EUI)



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

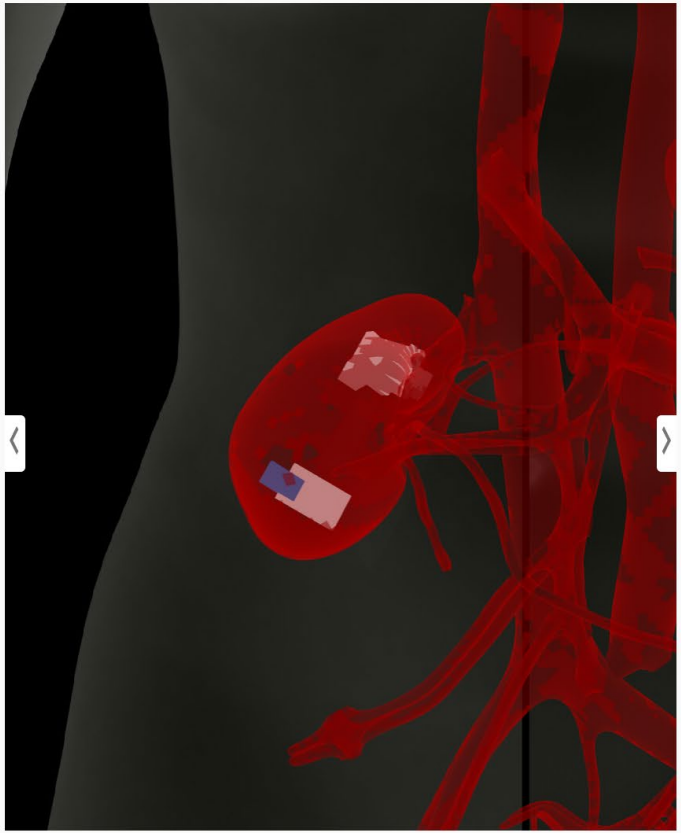
HuBMAP

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

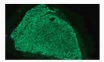


body

1 Centers
9 Donors
40 Samples

Male, Age 55, BMI 25.4
HBM695.RTLJ.484
TMC-Vanderbilt
13-16

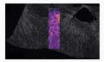
Male, Age 21, BMI 21.8
HBM634.MMGK.572
TMC-Vanderbilt
Age 21 , White Male, Trauma Patient



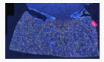
Female, Age 44, BMI 28.0
HBM457.NNQN.252
TMC-Vanderbilt
Age 44, white female.

Female, Age 44, BMI 28.0
HBM465.VKHL.532
TMC-Vanderbilt
Age 44, white female.

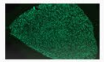
Male, Age 21, BMI 21.8
HBM693.HFFJ.752
TMC-Vanderbilt
Age 21 , White Male, Trauma Patient



Female, Age 58, BMI 23.0
HBM536.LDTZ.757
TMC-Vanderbilt
Age 58, White Female

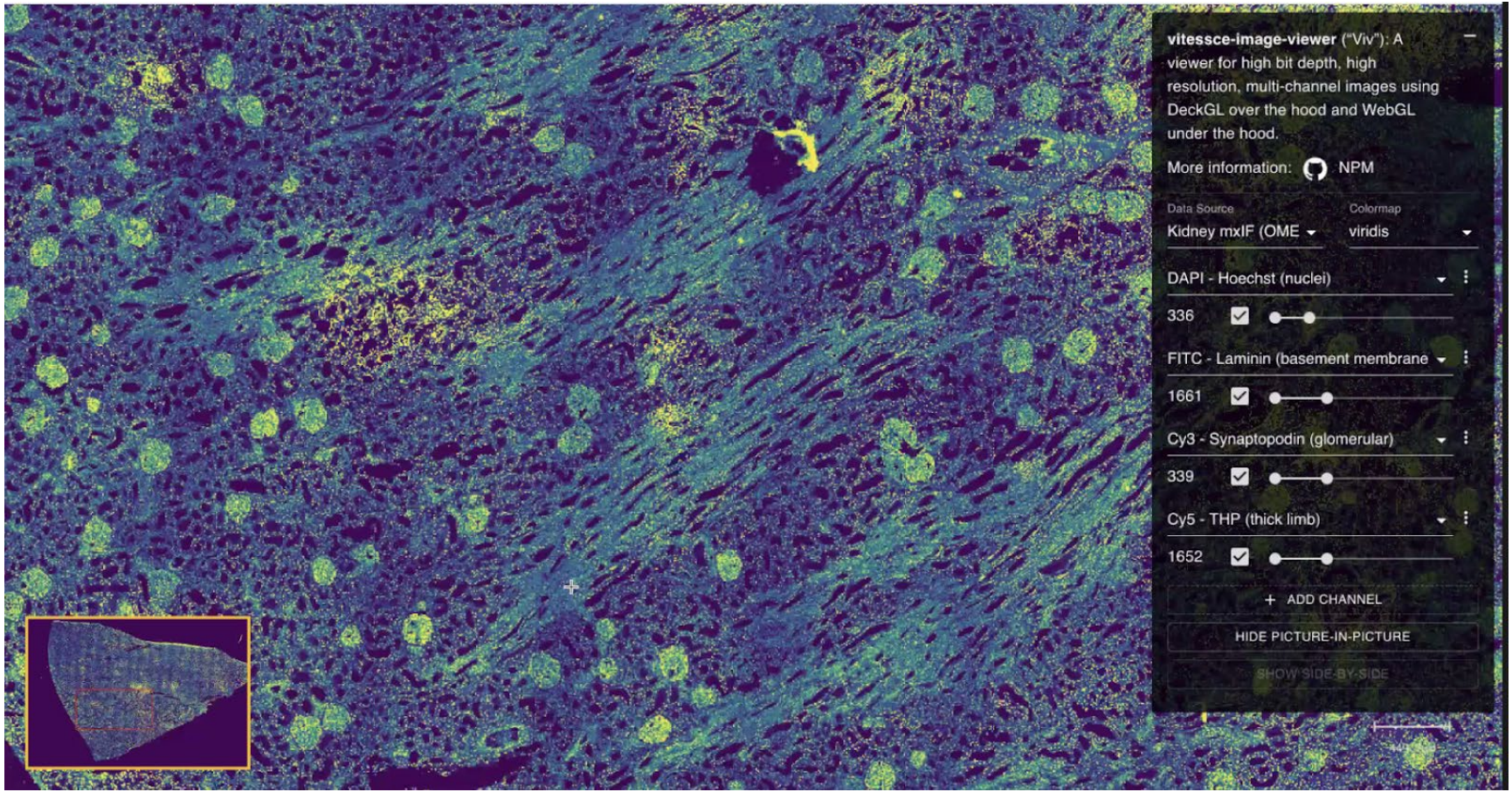


Male, Age 48, BMI 35.3
HBM334.GCCX.874
TMC-Vanderbilt
Age 48, White Male



Male, Age 31, BMI 32.6
HBM776.PKJF.786
TMC-Vanderbilt
Age 21, White Male

Female, Age 66, BMI 31.3
HBM284.TRCV.726



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

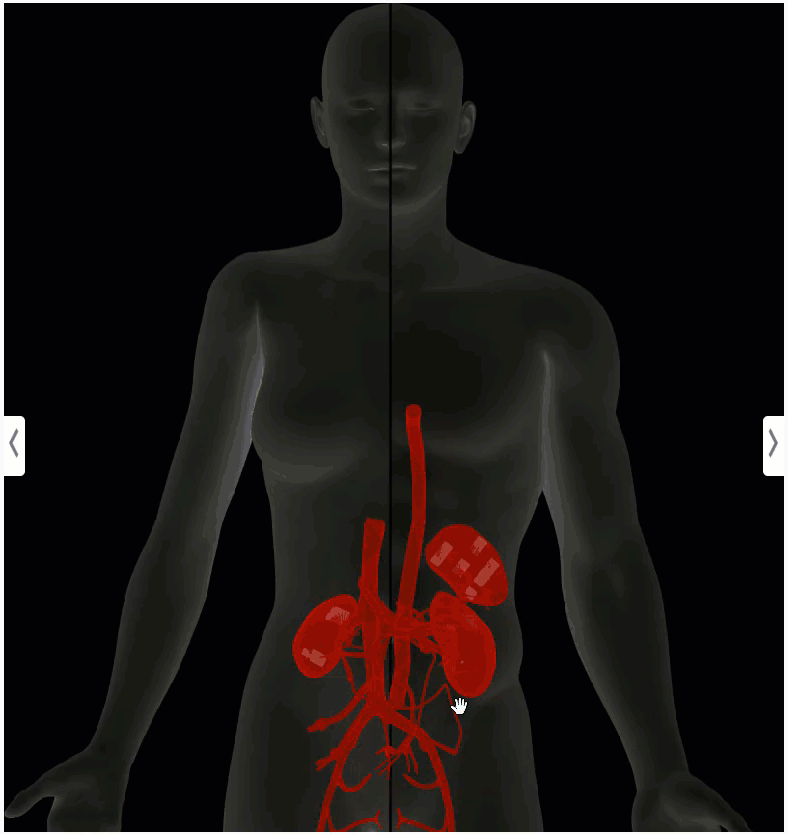
HuBMAP

Sex: Both Age: 1-110 BMI: 13-83

Login

Search ontology terms ...

- body
  - heart
  - lung
  - kidney
  - spleen
  - colon
  - small intestine
  - rectum

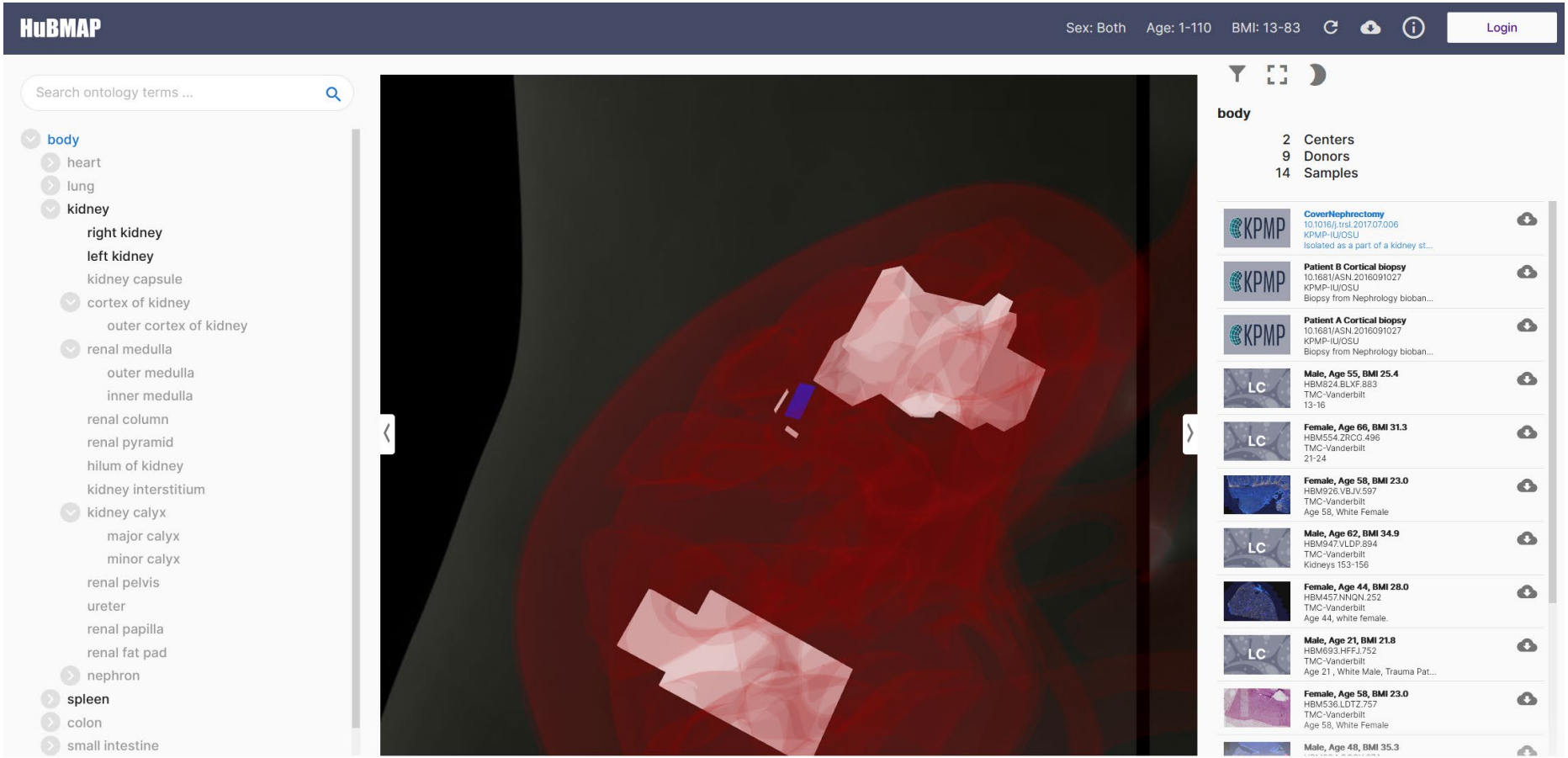


body

2 Centers
27 Donors
41 Samples

	Female, Age 58, BMI 23.0 HBM926.VBJV.597 TMC-Vanderbilt Age 58, White Female
LC	**Male, Age 46, BMI 22.3** HBM946.ZWHV.257 TMC-Vanderbilt 48-51
	Female, Age 76, BMI 37.5 HBM543.NGQC.475 TMC-Vanderbilt Age 76, white female.
LC	**Male, Age 55, BMI 30.0** HBM258.DDTW.423 TMC-Vanderbilt 52-55
LC	**Female, Age 38, BMI 42.3** HBM396.JRWZ.394 TMC-Vanderbilt 17-20
LC	**Male, Age 62, BMI 34.9** HBM947.VLDP.894 TMC-Vanderbilt Kidneys 153-156
LC	**Female, Age 44, BMI 48.2** HBM629.PPWR.872 TMC-Vanderbilt 25-28
CODEX	**Male, Age 18, BMI 27.1** HBM748.ZDKH.494 TMC-Florida Section is 400um from face edge ...

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



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

Cells of the adult human heart

https://www.nature.com/articles/s41586-020-2797-4

Monika Litviňuková, Carlos Talavera-López, [...] Sarah A. Teichmann

Nature (2020) | Cite this article

Published: 24 September 2020

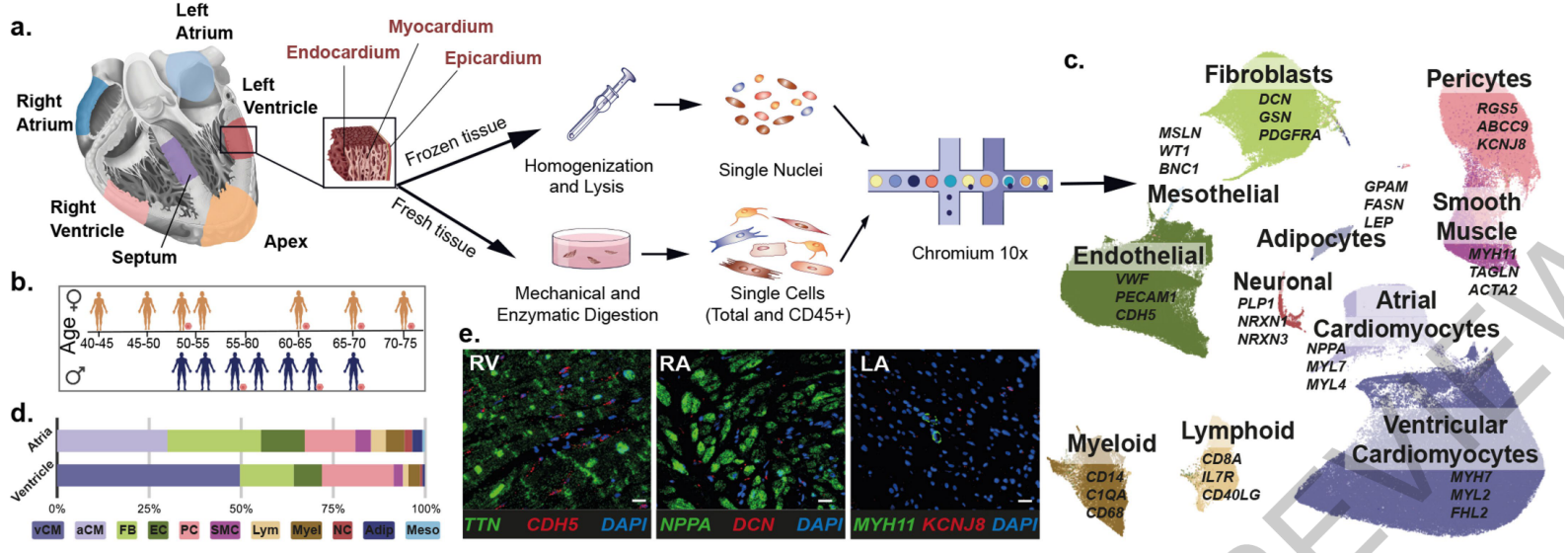


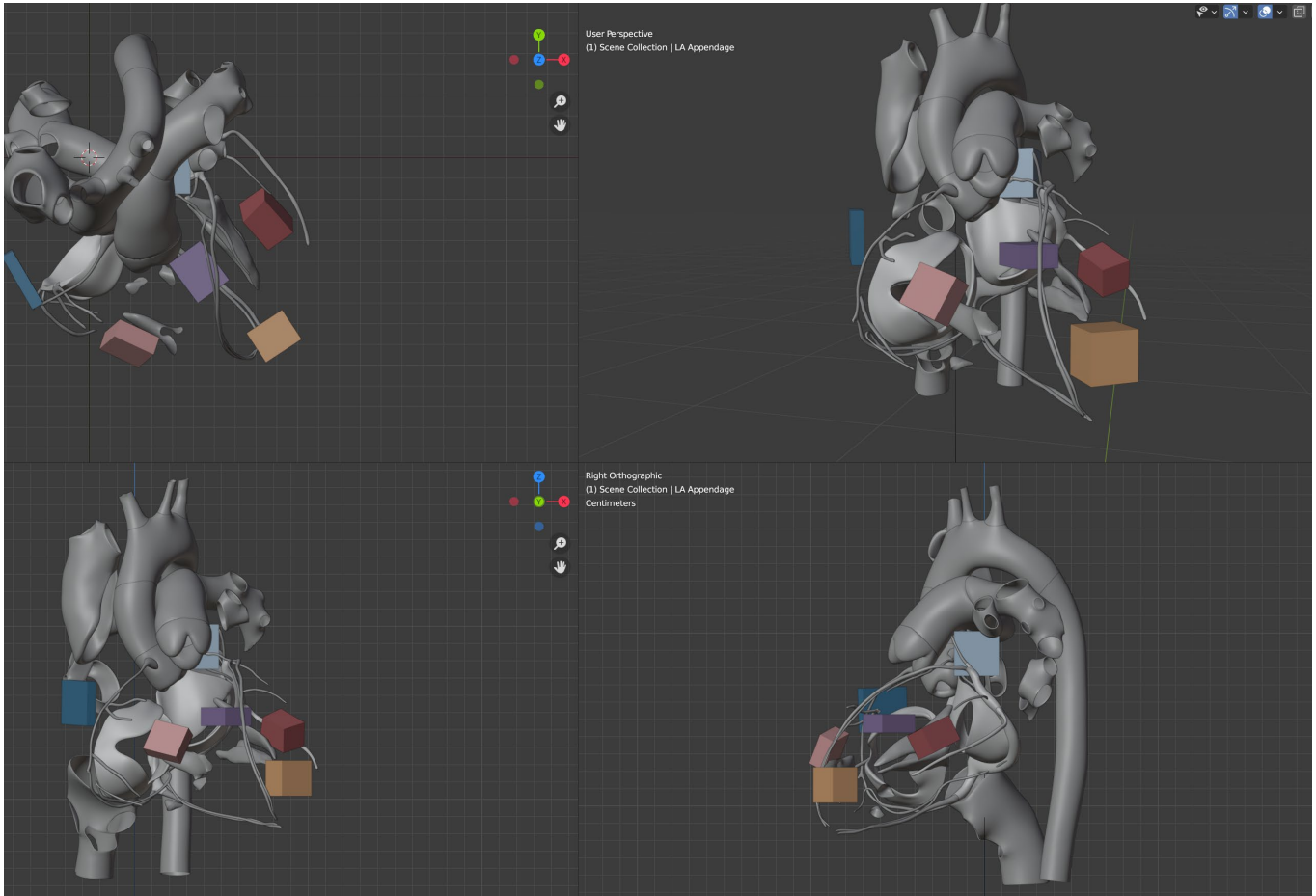
Fig. 1 | Cell composition of the adult human heart. a. Transmural samples were obtained from RA, LA, RV, LV, AX and SP from 14 individuals. Single nuclei (n = 14) and single cells (n = 7) were processed using Chromium 10X 3'DEG chemistry. **b.** Infographic shows donors (women, top; men, bottom), age, and contribution to cells and nuclei datasets (orange circle) (Data available in Supplementary Table 1) **c.** UMAP embedding of 487,106 cells and nuclei delineate 11 cardiac cell types and marker genes. **d.** Distribution of cell populations, identified from nuclei within atria (LA, RA) and ventricles (LV, AX, SP, RV) after subclustering analysis. Color code corresponds to **c** (Data available in Supplementary Table 2). **e.** Multiplexed smFISH of cell type-specific transcripts in RV (left): *TTN* (green, CM) and *CDH5* (red, EC) RA (middle): *NPPA* (green, aCM) and *DCN* (red, FB) and LA (right): *MYH11* (green, SMC) and *KCNJ8* (red, PC), nuclei are DAPI-stained (dark blue). Scale bar 20 µm. For details on statistics and reproducibility, please see **Methods**.

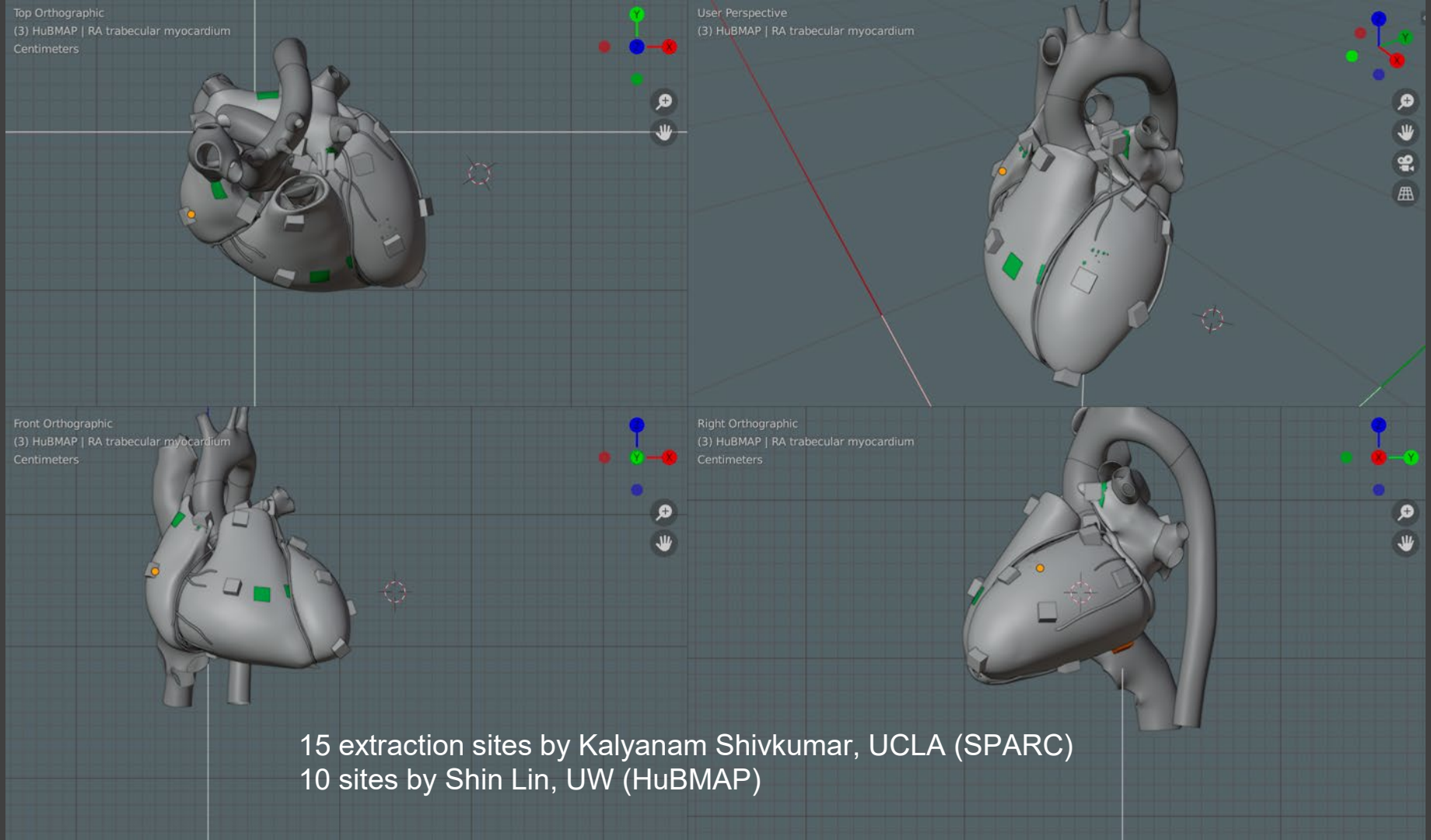
Cells of the adult human heart

Monika Litviňuková, Carlos Talavera-López, [...] Sarah A. Teichmann

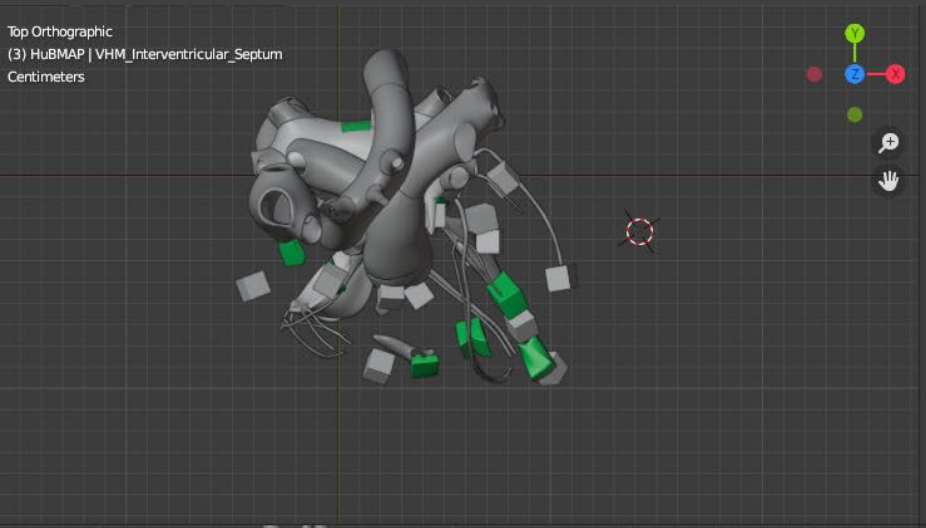
Nature (2020) | Cite this article

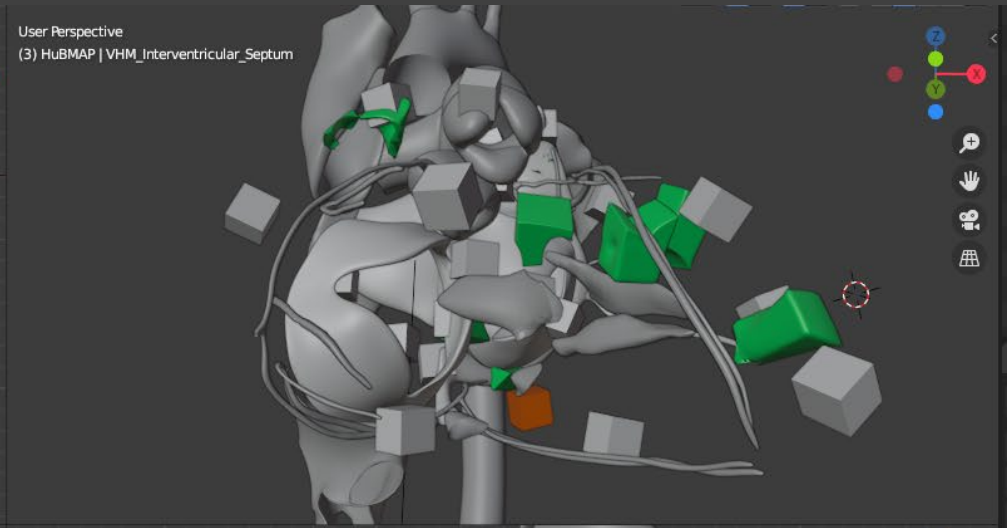
Published: 24 September 2020

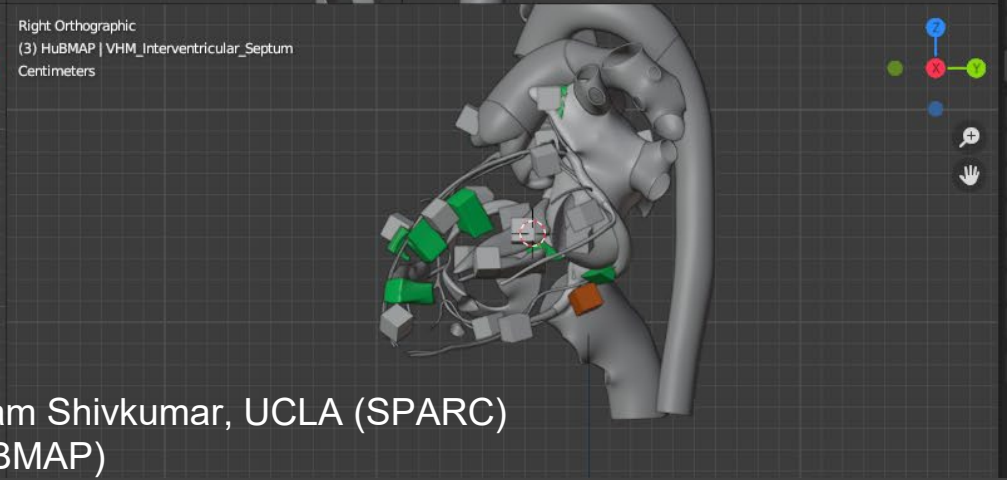




15 extraction sites by Kalyanam Shivkumar, UCLA (SPARC)
10 sites by Shin Lin, UW (HuBMAP)







15 extraction sites by Kalyanam Shivkumar, UCLA (SPARC)
10 sites by Shin Lin, UW (HuBMAP)

3D extraction sites do not restrict registration to specific regions, instead they provide “expert defined landmarks” to help guide tissue registration.

The extraction site objects are also used for automatic semantic annotation of tissue samples via collision detection during registration.

Shown here: Heart extraction sites

- 15 SPARC
- 10 HuBMAP
- 6 “Cells of the Adult Human Heart”

Human Reference Atlas CCF: Checklist

In support of Common Coordinate Framework (CCF) design (see CCF Portal):

1. Make sure the Anatomical Structures, Cell Types, and Biomarkers (ASCT+B) that you use/submit are listed in the ASCT+B tables. The tables are authored and reviewed by an international team of anatomists, pathologists, physicians, and other experts, see this SOP.
2. Spatially register all tissue samples using the CCF Registration User Interface (RUI) in the Ingest Portal. End of October 2020, kidney, spleen, heart, colon registration are supported. For other organs, see SOP.
3. After submitting data, review data in the CCF Exploration User Interface and make sure spatial, semantic, and other metadata are correct.
4. For functional tissue unit (FTU) segmentation, submit a list of FTUs for your organ(s) and make sure FTU names and all relevant cell types (CT) are captured in the ASCT+B table. Use assays/biomarkers (B) that make it possible to identify FTUs—initially manually, later automatically. Submit tissue with 1000 FTUs manually identified FTUs.
5. In support of the Vasculature-based CCF, provide cell segmentation data for blood vessels and different cell types.

For questions, email infoccf@indiana.edu.

HuBMAP Visible Human MOOC (VHMOOC)

Started Aug 4, 2020

To enroll, first **log in**. If you don't have an account, **create an IU Guest account**.

Register via:
https://tinyurl.com/vhmooc

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.

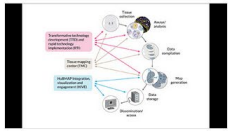
Andreas Bueckle, PhD Candidate in Information Science, performing research on information visualization, specifically virtual and augmented reality.

Length: 10 hours

Department: Cyberinfrastructure Network Science

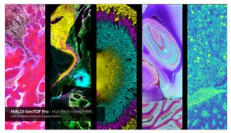
Credit: None

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



HuBMAP Overview

- Project Goals, Setup, and Ambitions

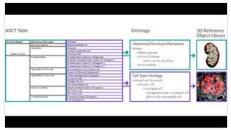


Tissue Data Acquisition and Analysis

- Behind the Scenes at Vanderbilt University

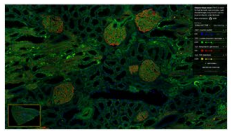
Biomolecular Data Harmonization

- An Introduction to Seurat



CCF Ontology, 3D Reference Objects, and User Interfaces

- Creating an Atlas of the Human Body



Portal Design and Usage

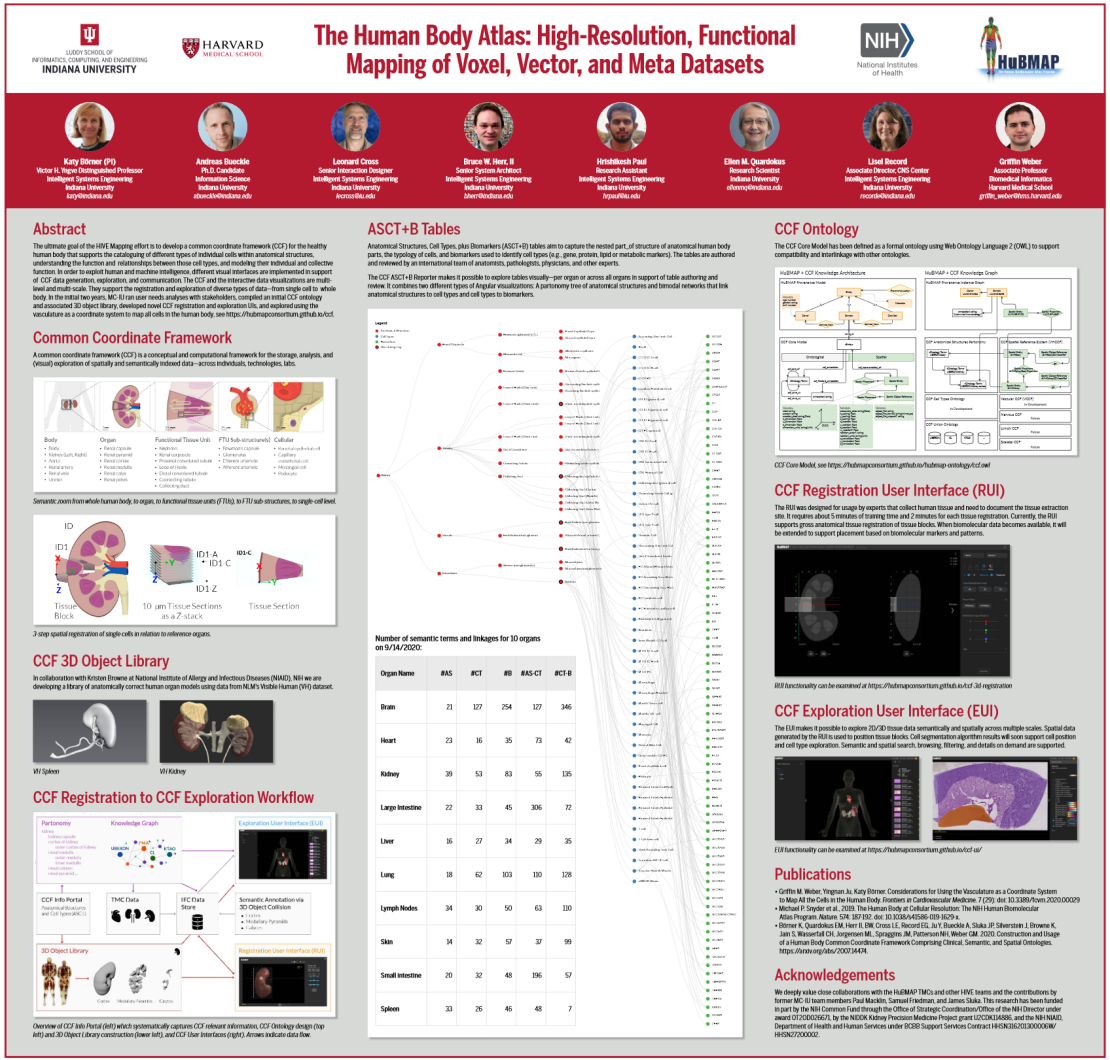
- Datasets and Software in the 1st HuBMAP Portal Release

Open Consent Your Data

- In Support of Research

Poster Session on Thursday, 11/19:

The Human Body Atlas: High-Resolution, Functional Mapping of Voxel, Vector, and Meta Datasets



The Human Body Atlas: High-Resolution, Functional Mapping of Voxel, Vector, and Meta Datasets

Katy Börner (PI), Victor A. Yngve Distinguished Professor, Intelligent Systems Engineering, Indiana University, katy@indiana.edu

Andreas Bueckle, Ph.D. Candidate, Information Science, Indiana University, abueckle@indiana.edu

Leonard Cross, Senior Interaction Designer, Intelligent Systems Engineering, Indiana University, lecross@iu.edu

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

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

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

Lisel Record, Associate Director, CNS Center, Intelligent Systems Engineering, Indiana University, record@indiana.edu

Griffin Weber, Associate Professor, Biomedical Informatics, Harvard Medical School, griffin_weber@hms.harvard.edu

Abstract

The ultimate goal of the HIVE Mapping effort is to develop a common coordinate framework (CCF) for the healthy human body that supports the cataloguing of different types of individual cells within anatomical structures, understanding the function and relationships between those cell types, and modeling their individual and collective function. In order to exploit human and machine intelligence, different visual interfaces are implemented in support of CCF data generation, exploration, and communication. The CCF and the interactive data visualizations are multi-level and multi-scale. They support the registration and exploration of diverse types of data—from single cell to whole body. In the initial two years, MC-IU ran user needs analyses with stakeholders, compiled an initial CCF ontology and associated 3D object library, developed novel CCF registration and exploration UIs, and explored using the vasculature as a coordinate system to map all cells in the human body. See https://hubmapconsortium.github.io/ccf.

Common Coordinate Framework

A common coordinate framework (CCF) is a conceptual and computational framework for the storage, analysis, and (visual) exploration of spatially and semantically indexed data—across individuals, technologies, labs.

Semantic zoom from whole human body, to organ, to functional tissue units (FTUs), to FTU sub-structures, to single-cell level.

3-step spatial registration of single cells in relation to reference organs.

CCF 3D Object Library

In collaboration with Kirsten Browne at National Institute of Allergy and Infectious Diseases (NIAID), NIH we are developing a library of anatomically correct human organ models using data from NLM's Visible Human (VH) dataset.

VH Spleen *VH Kidney*

CCF Registration to CCF Exploration Workflow

Overview of CCF Info Portal (left) which systematically captures CCF relevant information, CCF Ontology design (top left), and 3D Object Library construction (lower left), and CCF User Interfaces (right). Arrows indicate data flow.

ASCT+B Tables

Anatomical Structures, Cell Types, plus Biomarkers (ASCT+B) tables aim to capture the nested part_of structure of anatomical human body parts, the typology of cells, and biomarkers used to identify cell types (e.g., gene, protein, lipid or metabolic markers). The tables are authored and reviewed by an international team of anatomists, pathologists, physicians, and other experts.

The CCF ASCT+B Reporter makes it possible to explore tables visually—per organ or across all organs in support of table authoring and review. It combines two different types of Angular visualizations: A partonomy tree of anatomical structures and bimodal networks that link anatomical structures to cell types and cell types to biomarkers.

Number of semantic terms and linkages for 10 organs on 9/14/2020:

Organ Name	#AS	#CT	#B	#AS-CT	#CT-B
Brain	21	177	254	127	346
Heart	23	16	35	73	42
Kidney	39	53	83	55	135
Large Intestine	22	33	45	306	72
Liver	16	27	34	29	35
Lung	18	62	103	110	128
Lymph Nodes	34	30	50	63	110
Skin	14	32	57	37	99
Small Intestine	20	32	48	196	57
Spleen	33	26	46	48	7

CCF Ontology

The CCF Core Model has been defined as a formal ontology using Web Ontology Language 2 (OWL) to support compatibility and interlinkage with other ontologies.

CCF Core Model, see https://hubmapconsortium.github.io/hubmap-ontology/ccf.owl

CCF Registration User Interface (RUI)

The RUI was designed for usage by experts that collect human tissue and need to document the tissue extraction site. It requires about 5 minutes of training time and 2 minutes for each tissue registration. Currently, the RUI supports gross anatomical tissue registration of tissue blocks. When biomolecular data becomes available, it will be extended to support placement based on biomolecular markers and patterns.

RUI functionality can be examined at https://hubmapconsortium.github.io/ccf-3d-registration

CCF Exploration User Interface (EUI)

The EUI makes it possible to explore 2D/3D tissue data semantically and spatially across multiple scales. Spatial data generated by the RUI is used to position tissue blocks. Cell segmentation algorithm results will soon support cell position and cell type exploration. Semantic and spatial search, browsing, filtering, and details on demand are supported.

EUI functionality can be examined at https://hubmapconsortium.github.io/ccf-ui/

Publications

- Griffin M. Weber, Yingnan Ju, Katy Börner. Considerations for Using the Vasculature as a Coordinate System to Map All the Cells in the Human Body. *Frontiers in Cardiovascular Medicine*. 7 (79): doi: 10.3389/fcvm.2020.00029
- Michael P. Snyder et al., 2019. The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program. *Nature*. 574: 187-192. doi: 10.1038/s41586-019-1629-x.
- Börner K, Quardokus EM, Herr II, BW, Cross LE, Record EG, Ju Y, Bueckle A, Sluka JP, Silverstein J, Browne K, Jain S, Wasserfall CH, Jorgensen ML, Spraggins JM, Patterson NH, Weber GM. 2020. Construction and Usage of a Human Body Common Coordinate Framework Comprising Clinical, Semantic, and Spatial Ontologies. https://arxiv.org/abs/2007.14424.

Acknowledgements

We deeply value close collaborations with the HuBMAP TMCs and other HIVE teams and the contributions by former MC-IU team members Paul Macklin, Samuel Friedman, and James Sluka. This research has been funded in part by the NIH Common Fund through the Office of Strategic Coordination/Office of the NIH Director under award OT2OD026671, by the NIDDK Kidney Precision Medicine Project grant U2CDK114886, and the NIH NIAID, Department of Health and Human Services under BCBB Support Services Contract HHSN316201300006W/HHSN27200002.

Q&A