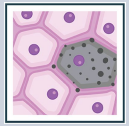


Human Organ Atlas

Hacking the Human Vasculature in 3D - A Kaggle Challenge

Presenter: Yashvardhan Jain
Research Software Engineer - Machine Learning
Cyberinfrastructure for Network Science Center
Department of Intelligent Systems Engineering
Indiana University
Bloomington, Indiana, USA

kaggle

CIFAR

ThermoFisher SCIENTIFIC

INDIANA UNIVERSITY

UCL

ROYAL ACADEMY OF ENGINEERING

CHAN ZUCKERBERG INITIATIVE

TECHPOINT

OUR ORGANIZERS

Katy Börner
Distinguished Professor of Engineering and Information Science
Indiana University Bloomington

Peter Lee
Professor of Materials Science
University College London

Claire Walsh
Senior Research Fellow
University College London

Yashvardhan Jain
Research Software Engineer - Machine Learning
Indiana University Bloomington

Lisel Record
CNS Associate Director
Indiana University Bloomington

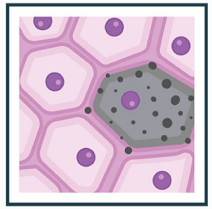
Nancy Ruschman
Project Manager
Indiana University Bloomington

Kaggle Team

Addison Howard
Head of Kaggle Competitions Program Management

Ashley Chow
Competitions Program Manager

Ryan Holbrook
Data Scientist



Human Organ Atlas

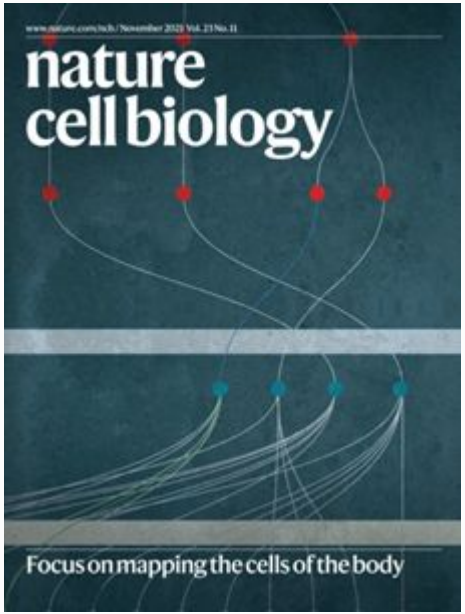
NIH
National Institutes of Health

Human Reference Atlas

The **Human Reference Atlas (HRA)** is a **comprehensive**, **high-resolution**, **three-dimensional** atlas of all the **cells** in the **healthy human body**. The Human Reference Atlas provides standard terminologies and data structures for describing specimens, biological structures, and spatial positions linked to existing ontologies.

The HRA

1. defines the 3D space and shape of anatomical structures and cell types that are of biomedical relevance plus the biomarkers used to characterize them. Anatomical structures, cell types and biomarkers are validated and represented in/added to ontologies (Uberon/FMA, CL, HGNC).
2. defines how new datasets can be mapped to the HRA, e.g., spatially using the Visible Human CCF or Vasculature CCF, via ASCT+B ontology terms/IDs, or via gene expression data as in Azimuth.
3. It is
    - authoritative (there exists expert agreement and/or experimental data),
    - computable (supports API queries, UIs),
    - published as LOD (connected to gene, disease, and other ontologies and data),
    - open (anyone can use the HRA data and code), and
    - continuously evolving (e.g., as new technologies become available).



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

Website: https://humanatlas.io/

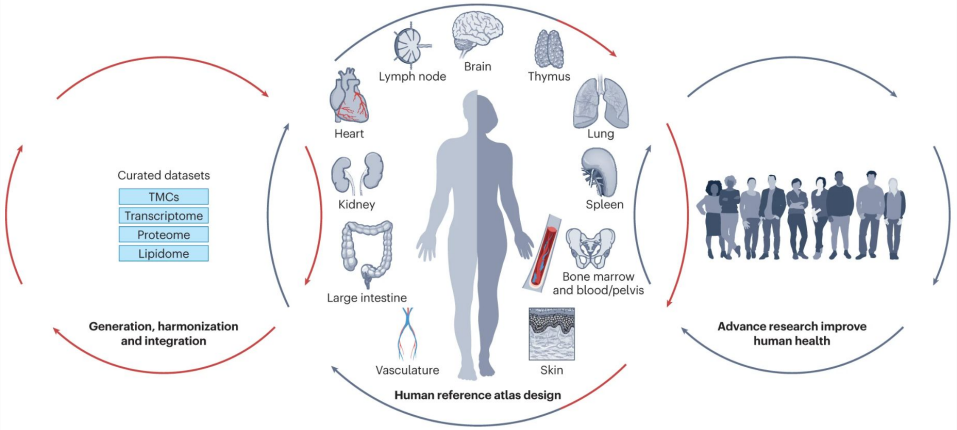
Cellular Senescence Network Program

Vision

The Common Fund's Cellular Senescence Network (SenNet) Program was established to comprehensively identify and characterize the differences in senescent cells across the body, across various states of human health, and across the lifespan.

Goals

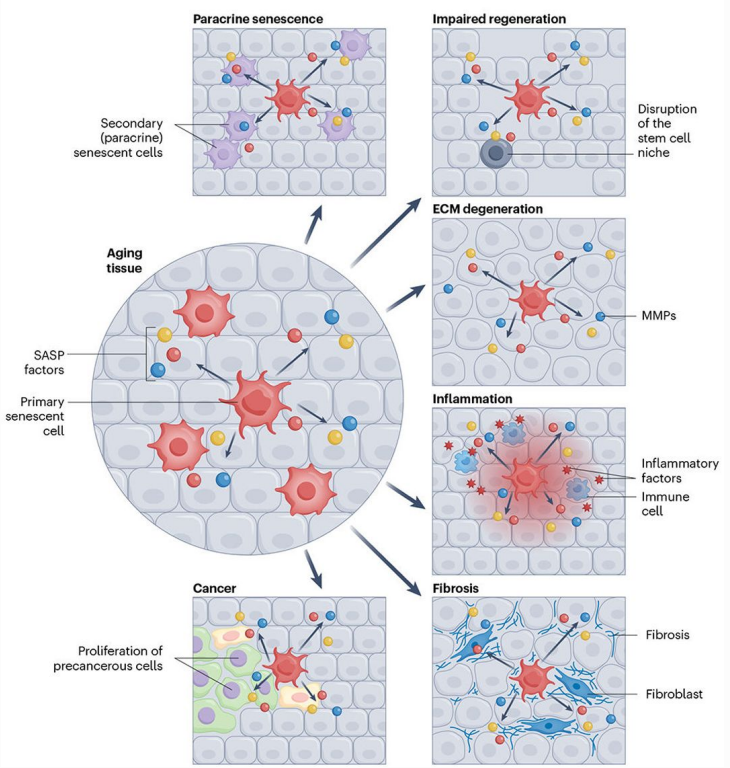
- Provide publicly accessible atlases of senescent cells, the differences among them, and the molecules they secrete, using data collected from multiple human and model organism tissues.
- To create tools & technologies and a reference atlas of senescent cells in healthy human tissues—accelerating the ability of biomedical researchers to develop therapeutics that target cellular senescence and improve human health.
- Unite cellular senescence researchers by developing common terms and classifications for senescent cells.



https://www.nature.com/articles/s43587-022-00326-5

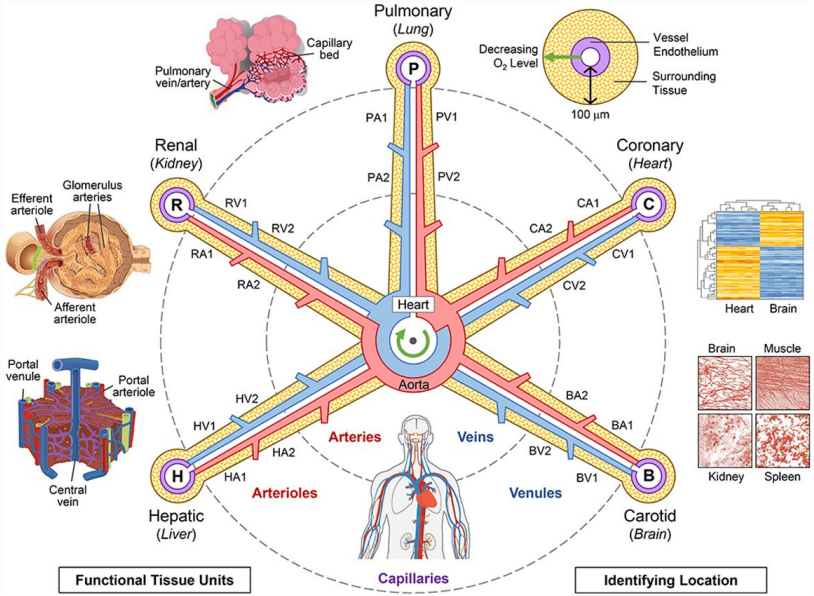
Exosomes

- Several teams within SenNet are studying **exosomes, membrane-bound extracellular vesicles used to transport diverse cargo to neighboring cells or even other organs.**
- The cargo could hurt (cause inflammation, senescence, cancer) or help (immune cells) other areas in the body and there is much active research on medical interventions using exosomes.
- Exosomes might explain comorbidities in humans or how healing one organ, e.g., a new kidney or heart, improves the function of many other organs.
- **Understanding the generation, transportation, and ingestion of exosomes requires data on the human blood vasculature—the main exosome transport system.**



https://www.nature.com/articles/s43587-023-00446-6

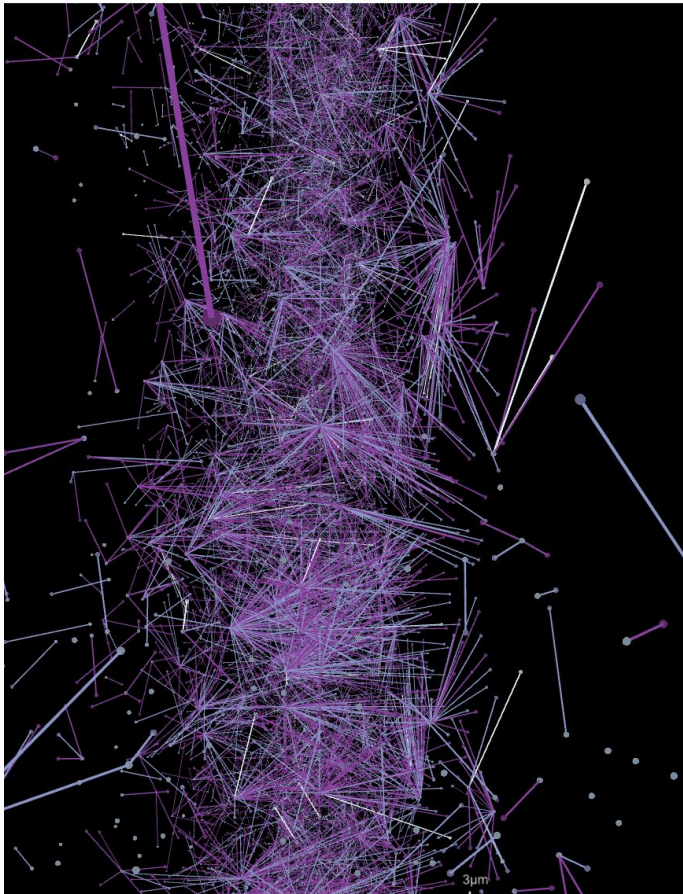
Vasculature Common Coordinate Framework



- The Vascular Common Coordinate Framework (VCCF) Database comprises 900+ vessels [1-2].
- Using experimental data, we can determine the pathway, throughput/vessel size, and 3D orientation in space of the human blood vasculature.
- Mapping the complex highway of blood flow will help examine exosome diffusion via the bloodstream to understand the role exosomes play in organ-to-organ communication or where exosomes might aggregate due to branches or tortuosity of the vessels.

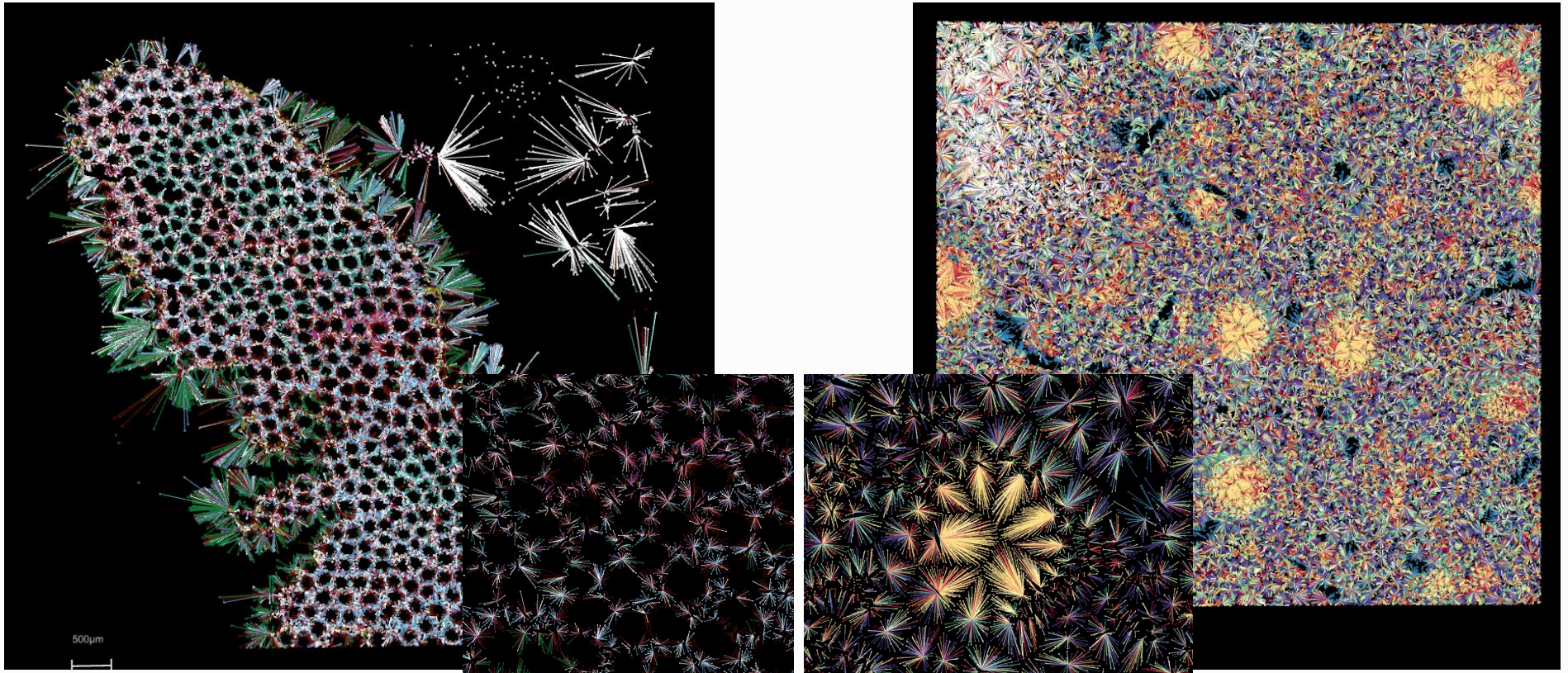
1. Weber GM, Ju Y, Börner K. Considerations for Using the Vasculature as a Coordinate System to Map All the Cells in the Human Body. *Front Cardiovasc Med.* 2020 Mar 13;7:29. https://www.frontiersin.org/articles/10.3389/fcvm.2020.00029/full
2. Database: Boppana A, et al. Anatomical structures, cell types, and biomarkers of the healthy human blood vasculature. *Sci Data*. 2023 Jul 19;10(1):452. https://www.nature.com/articles/s41597-023-02018-0

VCCF at the Single Cell Level



Skin Data from Ghose, S., Ju, Y., McDonough, E. *et al.* 3D reconstruction of skin and spatial mapping of immune cell density, vascular distance and effects of sun exposure and aging. *Commun Biol* 6, 718 (2023). https://doi.org/10.1038/s42003-023-04991-z

VCCF at the Single Cell Level



Intestine Data from Hickey, J.W., Becker, W.R., Nevins, S.A. *et al.* Organization of the human intestine at single-cell resolution. *Nature* 619, 572–584 (2023). https://doi.org/10.1038/s41586-023-05915-x

Spleen Data from Currlin, Seth, et al. "Immune, endothelial and neuronal network map in human lymph node and spleen." *bioRxiv* (2021): 2021-10.

Bridging Between Scales?

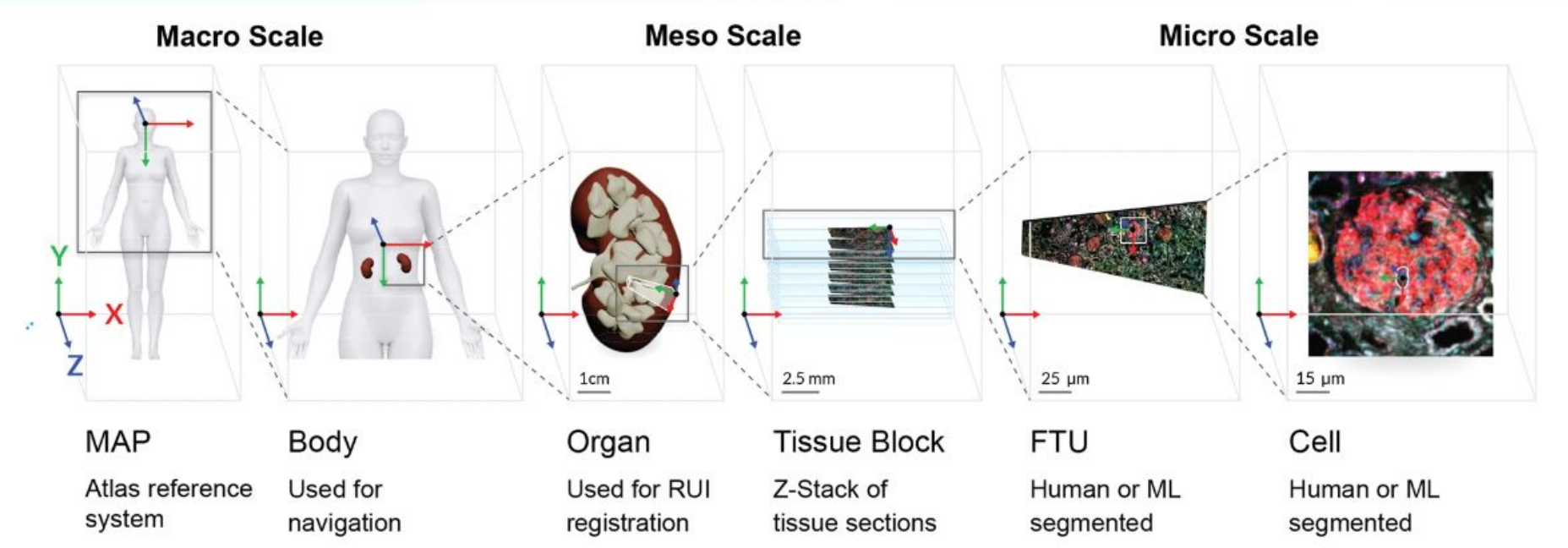
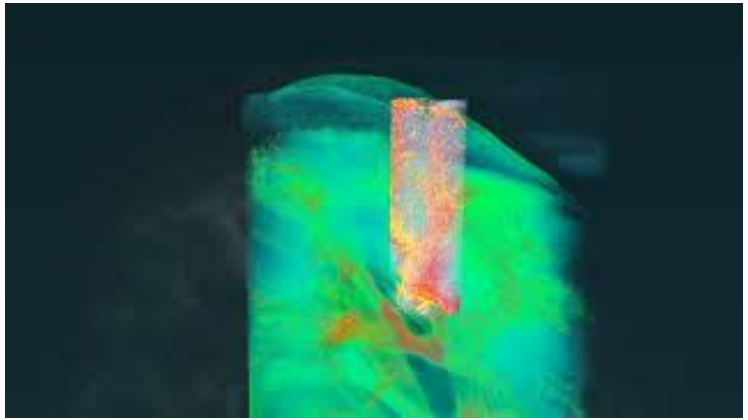
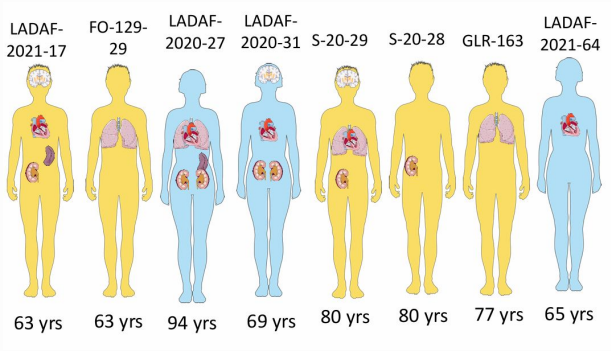


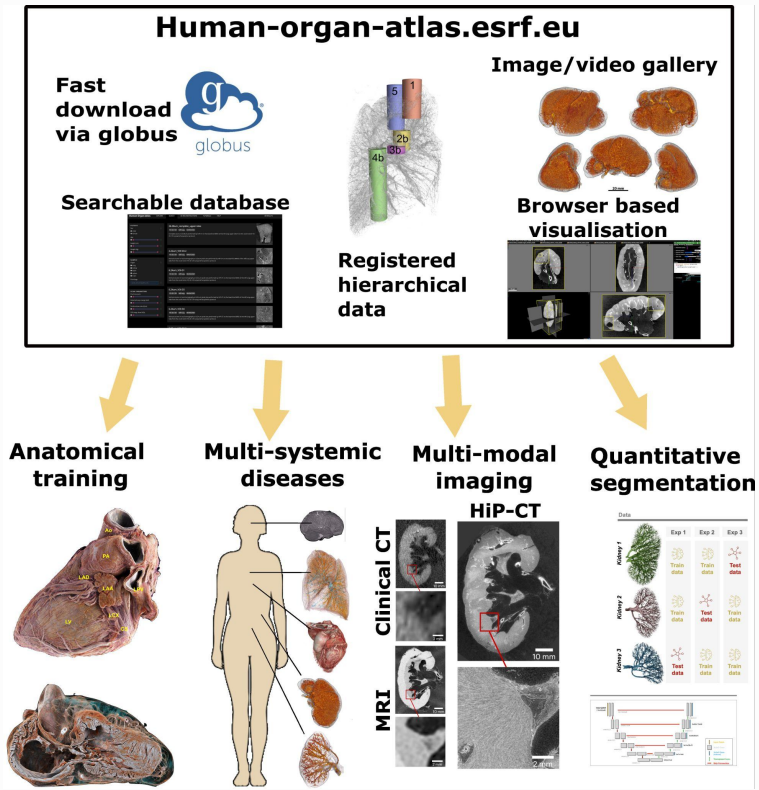
Image Source: Börner, K., Bueckle, A., Herr, B.W. *et al.* Tissue registration and exploration user interfaces in support of a human reference atlas. *Commun Biol* 5, 1369 (2022). https://doi.org/10.1038/s42003-022-03644-x

Human Organ Atlas



Hierarchical Phase-Contrast Tomography
Or HiP-CT: bit.ly/HiP-CT-videos



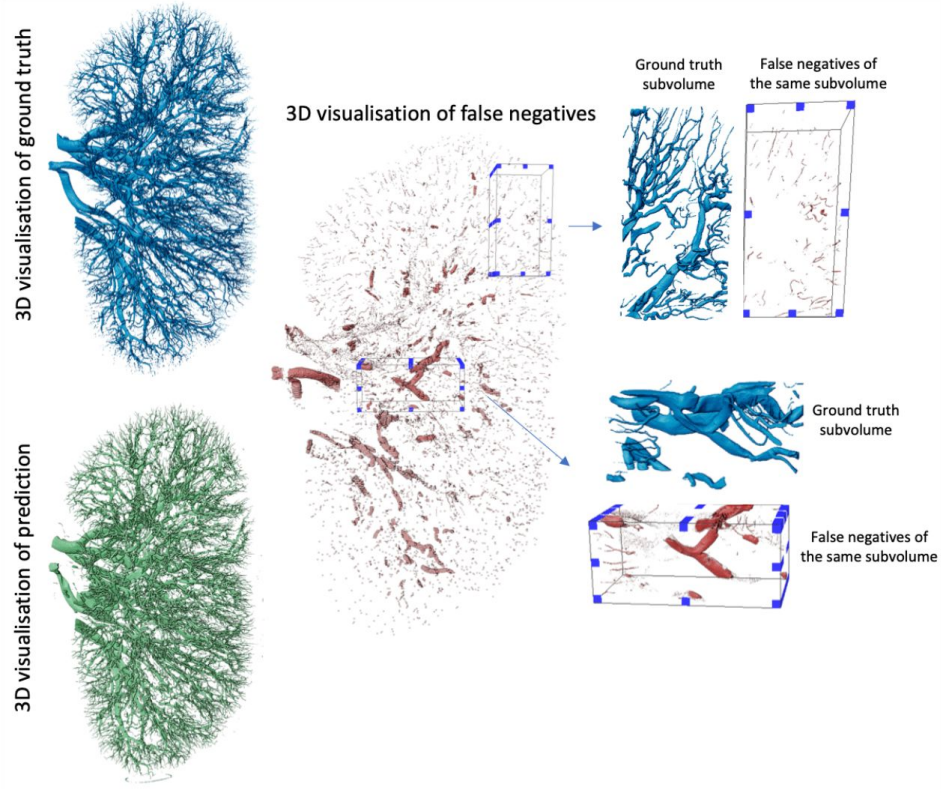


Database of human organs imaged with HiP-CT at multiple resolutions.

Available to download, view and utilise by the bioimaging community

See: https://mecheng.ucl.ac.uk/hip-ct/
Or Human-Organ-Atlas.esrf.eu

Issues with Vessel Segmentation



Vessel segmentation is challenging for a number of reasons:

1. Small vessels - **Imbalanced problem.**
2. **Low tolerance** for **connectivity error**.
3. **Collapse** or blood infilling of **vessels**
4. **Large variability** due to natural **human** anatomical **variation** and **post-mortem** variation
5. **Small training datasets** due to the **novelty** of the technique

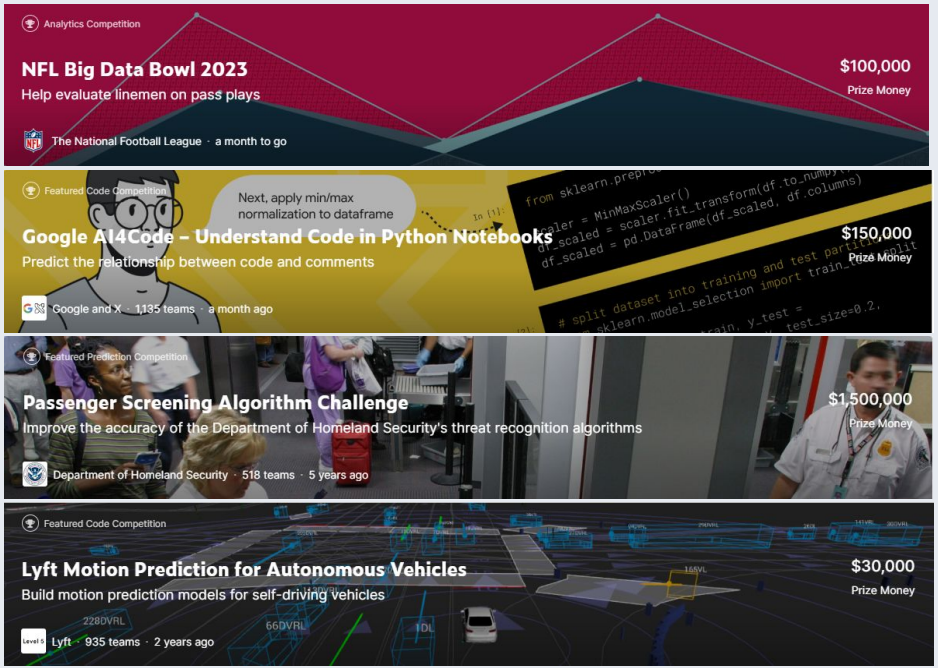
Yagis E, Aslani S, Jain Y, Zhou Y, Rahmani S, Brunet J, Bellier A, Werlein C, Ackermann M, Jonigk D, Tafforeau P. Deep Learning for Vascular Segmentation and Applications in Phase Contrast Tomography Imaging. arXiv preprint arXiv:2311.13319. 2023 Nov 22.

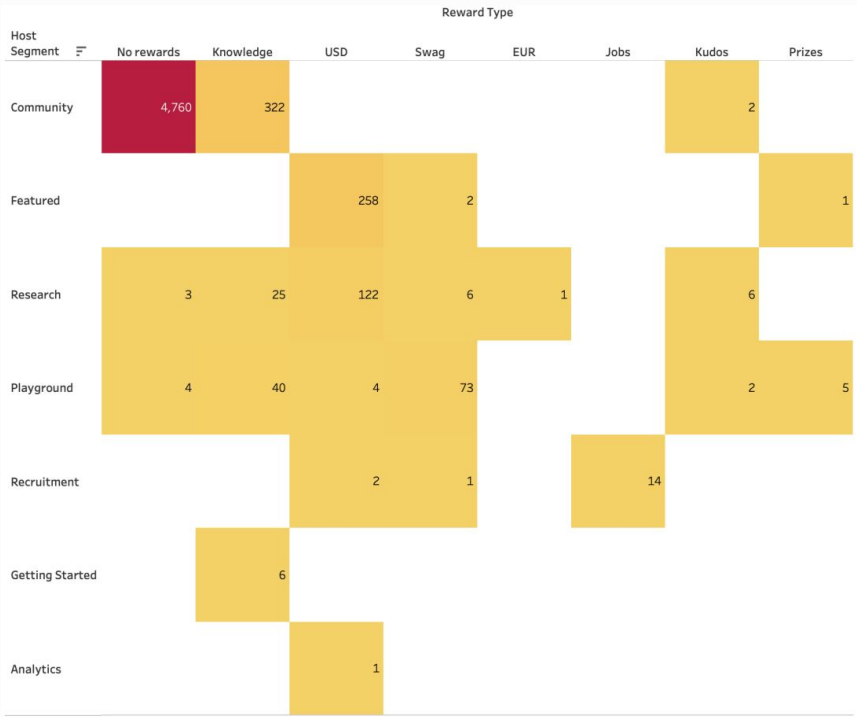
Kaggle - World’s largest machine learning community

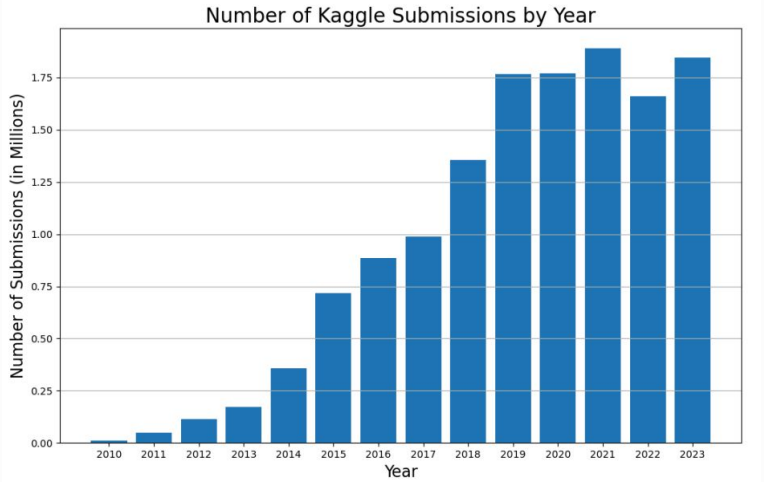
Kaggle is a data modeling and data analysis competition platform.

Businesses and researchers can publish data, and data scientists and machine learning practitioners can compete on the platform to produce the best models.

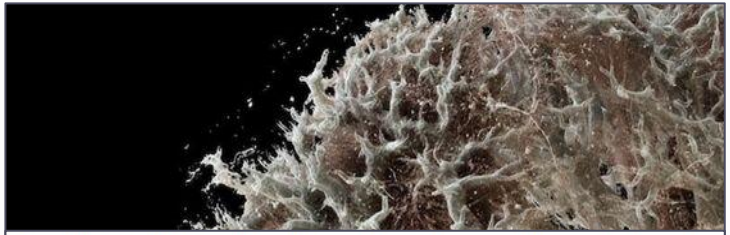
- Over 18 Million members
- Over 400,000 public notebooks
- Over 500 competitions
- Over 200,000 open datasets







Visualizations from work by Niteesha Jangam, Rasaghna Kuturu, Yamini Ane, Geetika Elaprolu, Pranay Reddy Gundala, Kabir Chaturvedi, Yeon-Soo Chung, Bhoomika Pathapati, Monisha Patro, Ishika Thakur, Venkata Viswanath Chittilla, Dwarakamai Mannemuddu, Srikeerthana Reddy Bandi, Poojitha Mathi, Sai Sumanth Muvva at Indiana University.

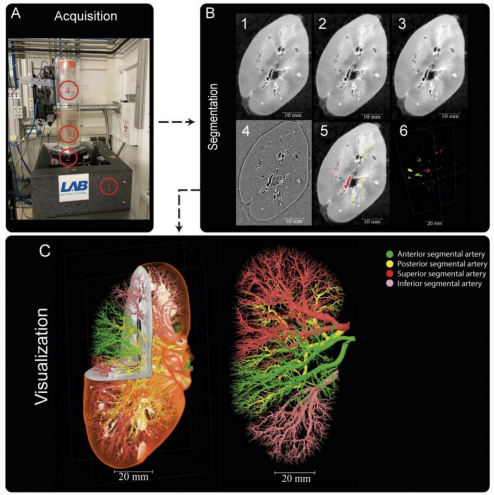


SenNet + HOA - Hacking the Human Vasculature in 3D

Segment vasculature in 3D scans of human kidney

Start	Close
Nov 7, 2023	Feb 6, 2024

Merger & Entry

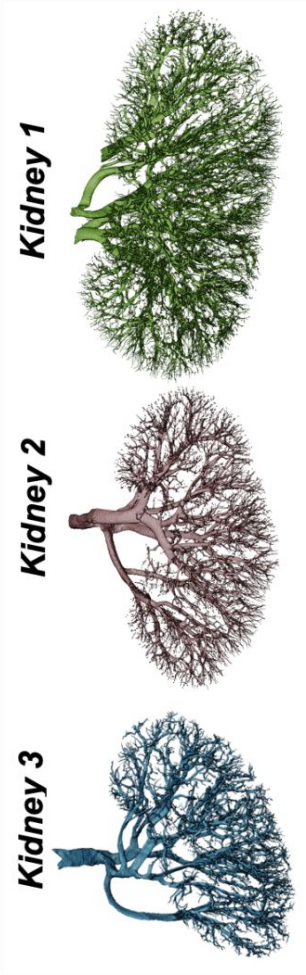


The data

- **5** Whole kidneys
- Images at **~50um/voxel**
- **3 Training** kidneys
- **2 Test** Kidneys.
- Additional data at ~5um/voxel.

- All data **manually segmented**
- **Triple validated** by expert annotators
- over **600 hours** of expert segmentation time.
- **Sparse labels** for some datasets,.

- **Baseline** model testing using **nn-Unet** provided.



Yagis E, Aslani S, Jain Y, Zhou Y, Rahmani S, Brunet J, Bellier A, Werlein C, Ackermann M, Jonigk D, Tafforeau P. Deep Learning for Vascular Segmentation and Applications in Phase Contrast Tomography Imaging. arXiv preprint arXiv:2311.13319. 2023 Nov 22.

Winning Teams

#	△	Team	Members	Score	Entries
1	—	Clevert		0.774105	278
2	▲ 1022	ryo		0.755959	33
3	▲ 568	ForcewithMe		0.727912	114
4	▲ 1	Igor Krashenyi		0.712188	251
5	▲ 5	Ivan Panshin		0.691797	108

kaggle CIFAR ThermoFisher SCIENTIFIC

Sponsoring a total of **$80,000** in prize money, making this competition possible and contributing to major progress in scientific research.

Key Solutions

- **Intensity-based** and **rotation augmentations.**
- **Boundary** focused **loss function** e.g. Marching cube loss
- **Post processing**
- Refined **pseudo-labeling** strategies
- **Ensembles** of **2D** and **2.5D** models generally performed better than 3D models.

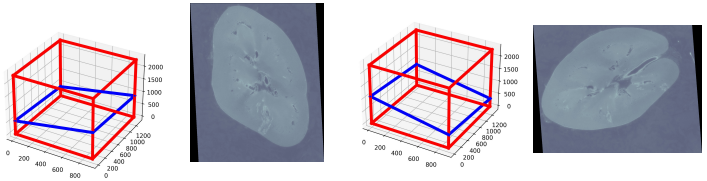
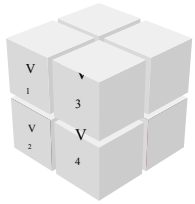


Fig. adapted from Team 1 Clevert



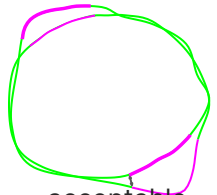


Fig. adapted from Team 1 Clevert

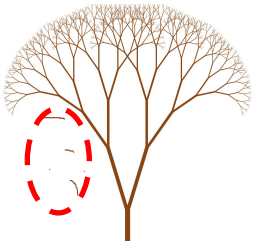
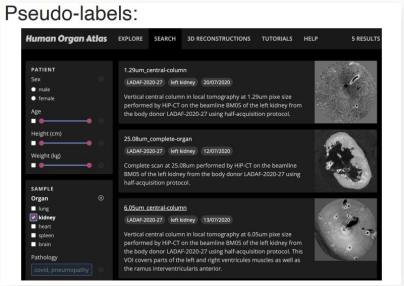


Fig. adapted from Team 2 ryo

Pseudo-labels:



Power of the Kaggle Community

1,401 participants
1,149 teams

6,128 individuals joined

204 first time Kagglers (including 5 users in the Top 100!)

32,391 submissions

500 Public Notebooks

756 Discussion Comments
200 Discussion Forum topics

A truly global challenge!
78 Countries

Acknowledgments

OUR ORGANIZERS

Katy Börner
Distinguished Professor of Engineering and Information Science
Indiana University Bloomington

Peter Lee
Professor of Materials Science
University College London

Claire Walsh
Senior Research Fellow
University College London

Yashvardhan Jain
Research Software Engineer - Machine Learning
Indiana University Bloomington

Lisel Record
CNS Associate Director
Indiana University Bloomington

Nancy Ruschman
Project Manager
Indiana University Bloomington

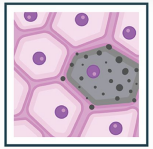
Kaggle Team

Addison Howard
Head of Kaggle Competitions Program Management

Ashley Chow
Competitions Program Manager

Ryan Holbrook
Data Scientist

Thanks go to all 1,401 **Kagglers** who participated along with the larger Kaggle community!



Human Organ Atlas

NIH National Institutes of Health

ROYAL ACADEMY OF ENGINEERING

CHAN ZUCKERBERG INITIATIVE

TECHPOINT

INDIANA UNIVERSITY

UCL

kaggle

CIFAR

ThermoFisher SCIENTIFIC

https://www.kaggle.com/competitions/blood-vessel-segmentation

Thank You!

Contact: yashjain@iu.edu