# Team 5

Visualizing Sample and Ontology linkages between **HuBMAP** and **SPARC**

# Contributing Team Members

## Bruce W. Herr II

Senior System Architect / Project Manager - Cyberinfrastructure for Network Science Center at Indiana University

## Samuel O’Blenes

Supervisor research web development at HealthPartners institute

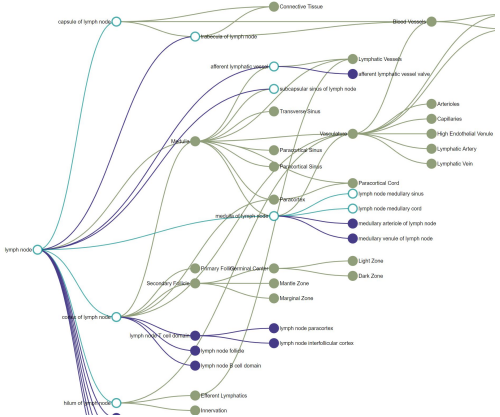
# Motivation

HuBMAP and SPARC have a rich reference system, both ontologically and spatially, tied to massive amounts of tissue sample data. Both systems are growing organically and are in danger of being incompatible. To enable future cross-pollination and compatibility, we have implemented a system to assist with mapping and visualization of the overlaps between the systems.

# HuBMAP & SPARC Linkages

The product has four modules to facilitate cross-consortium data compatibility:

- Inter-system ontological network visualization
- Inter-consortium dataset comparison
- SPARC-enabled CCF-EUI visualizer
- ASCT+B SPARC data export facility

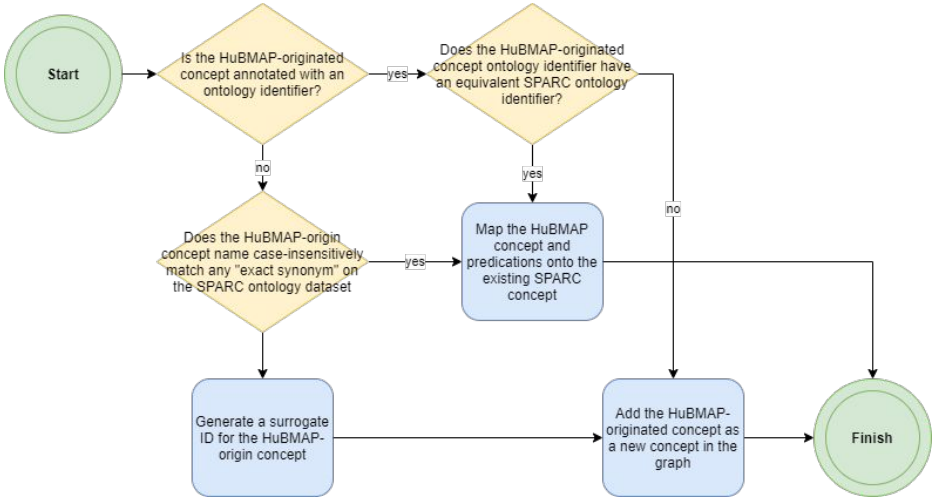


# Data sources

| Data type | Source | Method |
|---|---|---|
| HuBMAP ontology concepts | CCF ASCT+B Reporter ASCT+B API | The API references production versions of the HuBMAP CCF ASCT+B tables where concepts are curated |
| SPARC AS ontology concepts | FDI Lab SciCrunch/SciGraph API | Organ anatomical structure concept datasets are extracted via neighbor search referencing inbound PART_OF predicate up to 10 degrees |
| SPARC CT ontology concepts | FDI Lab SciCrunch/SciGraph API | Cell Type concepts are extracted via cypher query on UBERON-to-Cell Ontology single-hop linkages across all predicates |
| HuBMAP dataset metadata | HuBMAP Elastic Search API | Dataset metadata is ingested directly from the endpoint, filtered to CCF-registered datasets |
| SPARC dataset metadata | SPARC Elastic Search API | Dataset metadata is filtered to human-species subset |

# Ontology concept merging

- First pass at merging is based on ontology identifier comparison
  (UBERON ID in most cases)
- Disambiguation where HuBMAP concepts not annotated with an ontological identifier:
  - The concept name is compared to exact synonyms among participating SPARC nodes
  - If no synonym matches exist in SPARC, the concept is treated as novel and added to the graph



# Pre-existing resources

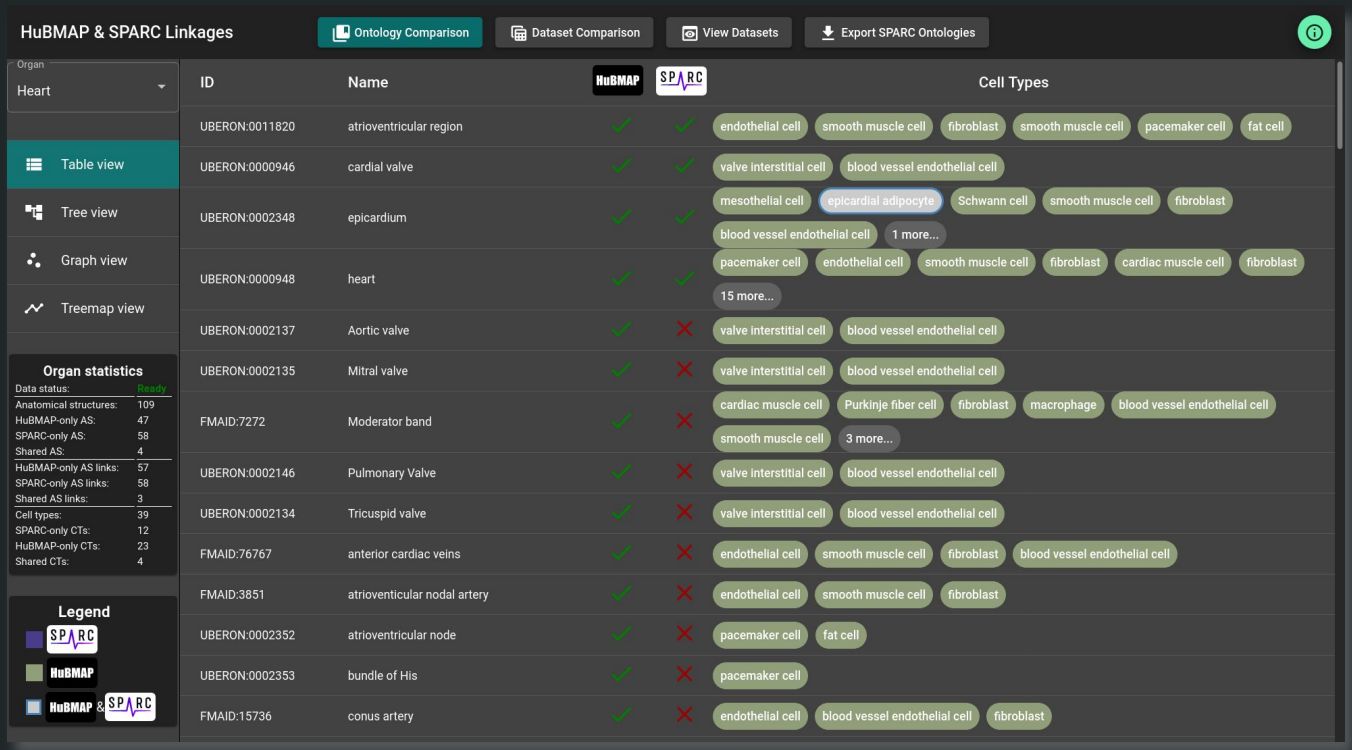
The following infrastructure items that will appear in today's demonstration were developed prior to this team's participation in the SPARC FAIR Codeathon:

- HuBMAP CCF Explorer User Interface (CCF-EUI)
- HuBMAP CCF ASCT+B Reporter

However, note that SPARC data ETL into formats compatible with these two HuBMAP facilities is new and was not available prior to our submission for the Codeathon

All other software that we will demonstrate today (ie. the HuBMAP and SPARC Linkages application) is new development for the Codeathon

# Product Demonstration



# Implementation highlights

- Web application (Angular) plus API microservice (Express) architecture
- Use of web-components to embed the CCF Exploration User Interface
- CI/CD workflow built in
- Live, always up-to-date resources
- Code documentation, UI hints/descriptions, and repository docs
- Project management with GitHub Issues and Project Boards

# Impact - Key Takeaways

- A great start for HuBMAP and SPARC to systematically compare and share data and ontologies
- A dashboard to track progress
- An initial method for adding SPARC scaffolds to the HuBMAP CCF
- A first implementation for integrating SPARC datasets into the CCF
- A path for SPARC to be able to contribute to the active ontology development efforts around ASCT+B working group (100+ individual members)

# Future Directions

- User studies to determine what metrics, visualizations, and features would be useful to SPARC & HuBMAP stakeholders
- Add more visualizations and aggregations to the dashboard
- Finalize SPARC scaffold to CCF Reference Organ interoperability
- Better placement of SPARC datasets in the CCF reference system
- Go the other direction (CCF -> SPARC)

# Thank you!

**Any questions?**