

The screenshot displays the Protégé ontology editor interface. The top tabs include 'Active Ontology', 'Entities', 'Individuals by class', and 'DL Query'. The left pane shows a class hierarchy with 'whole-cell patch-clamp voltage clamp assay' selected. The right pane shows the 'Class Annotations' for this class, including 'definition source' (Dr. Brenda Farrell), 'device' (patch clamp device and patch pipette), 'editor preferred label' (whole-cell patch-clamp voltage clamp assay), and 'term editor' (Jason Bengtson). Below the annotations, the 'Description' is shown as a logical definition:

```

achieves_planned_objective some 'assay objective'
has_specified_output some
  ('measurement datum'
   and ('is about' some 'ion channel activity'))
(realizes some
  (function
   and ('inherits in' some 'voltage clamp device')))
and (has specified input some 'voltage clamp device')

```

COLLABORATING WITH DATA: BUILDING A PARTNERSHIP WITH AN ELECTROPHYSIOLOGY RESEARCHER

Jason Bengtson Kansas State University

Dr. Brenda Farrell Baylor College of Medicine

WHY DO WE NEED TO COLLABORATE?

- Information and Data Science *add value*
 - Collaboration is increasingly important to academia and industry
 - Through collaboration, efforts become magnified
 - Collaboration allows different skill sets and expertise to complement each other
-

THIS PROJECT

- Electrophysiology data in Matlab data files
 - Adding metadata to make it as FAIR compliant as possible
 - Findable
 - Accessible
 - Interoperable
 - Reusable
 - Converting it into a less proprietary format
-

COMING TOGETHER

- Dr. Farrell approached the Texas Medical Center Library
 - As the Innovation Architect, I was asked to work on the project
 - I continued with the project when I left TMC for KSU
-

REFINING OUR APPROACH

- The Data Architect developed the initial data framework
 - We shared concerns about applying adequate metadata
 - We decided to base our approach on an ontology (OBI)
 - HDF5 files provided a group structure which could mirror ontological classes
 - We decided to map an ontology to HDF5 to contain data points
-

REFINING OUR APPROACH PT. 2

- . No ontology met our needs completely
 - . We had to graft other classes into OBI
 - . HDF5's attribute model allowed us to add definitions to groups
 - . In modifying the ontology we followed Minimum Information to Reference an External Ontology Term (MIREOT) guidelines
 - . The Researcher, given her extensive subject knowledge, worked with the OBI curators to add/refine classes
-

WORKING TOGETHER

- Distance and other commitments have created obstacles to collaboration
 - We used a variety of collaboration tools
 - Zoom for meetings
 - Dropbox
 - Webprotege
-

PRESENTING

- . This is the third conference that we have presented to
 - . The first was Association for Research in Otolaryngology 2017 in Baltimore, Maryland
 - . The second was ICBO 2018 in Corvallis, Oregon
 - . We collaborate on content
 - . One of us travels to the conference to present (the Researcher at the first two)
-

NEXT STEPS

We are in the final stages of preparing a paper for publication

Describes our data format

Compares it to alternatives

Within a month or two we will make the final submission of the Researcher's data

WHAT WE LEARNED ABOUT DATA CURATION

- a. The Researcher realized how woefully inadequate her former descriptions were and became educated on curation procedures.
 - b. The Data Architect became aware of the latticework of intricate and often vast frameworks that describe a researcher's data structures.
 - c. Both of us realized there is a need to provide tools and procedures to facilitate the process and ensure research data retains its value.
-

WHAT WE LEARNED ABOUT PARTNERSHIPS

- a. Be respectful of each other's time and expertise,
 - b. Provide a plausible time frame to accomplish each goal.
 - c. Fix the problems as they arise.
 - d. Acknowledge when someone is more proficient at one task over another, and let them when possible perform the task.
-

Our Thanks

This project was funded by supplemental Informationist grant NIDCD R01DC00354-SI awarded by NIH-NLM

Questions

