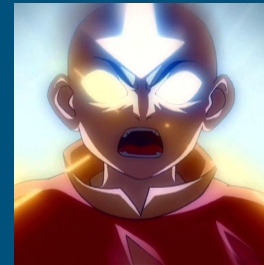# Sprint 2 Proposal:
13 November 2017

AVATR

AVATR will focus on LIMS for Sprint 2

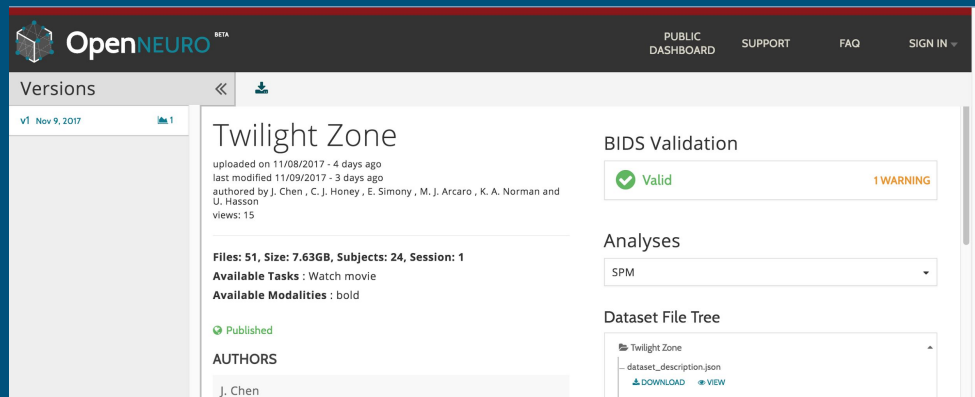| Sprint 2	12/20	- Refine Annotation workflow - What worked in sprint 1? What didn't? - Develop MVP for LIMS with the following features: - Dataset registration - Dataset queries - Tentative: Analyses and metadata logging - User handling - Good Documentation

Data Registration: Where is our data being stored?

- It turns out we don't really know. Data can be stored in the BOSS, an S3 Bucket, or even a single person's laptop.
- By registering datasets in our LIMS, we will know exactly where our datasets are living and how we can access them.

For the MVP, Users will manually register their data into LIMS.

- **Definition of Done**: Web form that handles dataset information and stores data specifications correctly in MongoDB.
- Specifications can include: keywords, owner, link, permissions, etc.
- Nice-to-haves: editing/deleting registered data

Dataset Query: Users should be able to search our LIMS for particular datasets.

- **Definition of Done**: MVP search tool. User provides keywords and we return datasets that best match those keywords.
- Nice to haves: search filters: specific data type, data owner, etc.
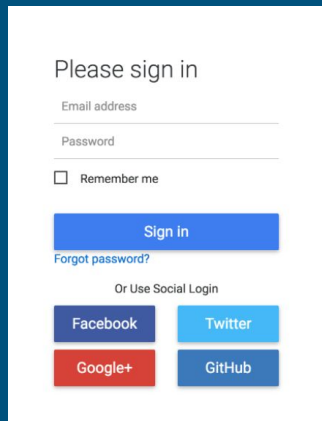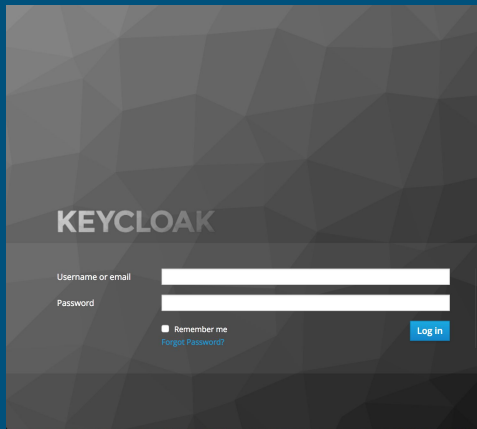
Logging: How to keep track of pipeline results and metadata?

- We want to know and look at specific pipeline results. Thus, would be nice to have a log of who ran the pipeline with what parameters, and what the results were.
- Problems with feasibility:
  - Users can't manually handle logging (would be very painful).
  - Every pipeline can be run in a different setting. Difficult to have a generalized logging tool.
- Definition of Done: Do more research on logging tools. If feasible, work with a specific pipeline and try to get logging working for it.

User Handling in the LIMS

- This is mostly just for handling data permissions and authorization.
- Potentially not as important for MVP but we definitely need this for our LIMS eventually.
- Definition of Done: User login and validation via OAuth or Keycloak

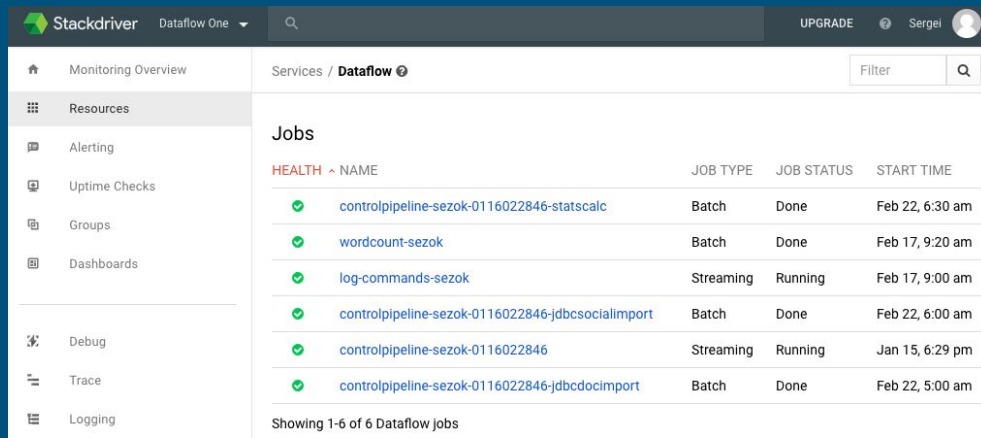# What's left…

- Good documentation
- Good frontend

Beginning Steps

- Wireframe of how our LIMS will work end-to-end.
- Database schemas for how information should be organized in LIMS.
- Tech Evaluations on current landscape:
  - Data registration: OpenNeuro, ndontology, render (mongodb schemas)
  - Data query: ElasticSearch?
  - Logging: Stackdriver