



# MODULAR APPLICATIONS FOR DATA PIPELINE AUTOMATION

Ashton Teng  
September 7, 2018

“

*The Big Picture:*

*To build software and data pipelines that are reliable, flexible, and automated in order to best serve iterative biological experiments that produce large amounts of data.*



Don't Repeat Yourself (DRY)  
Modularity

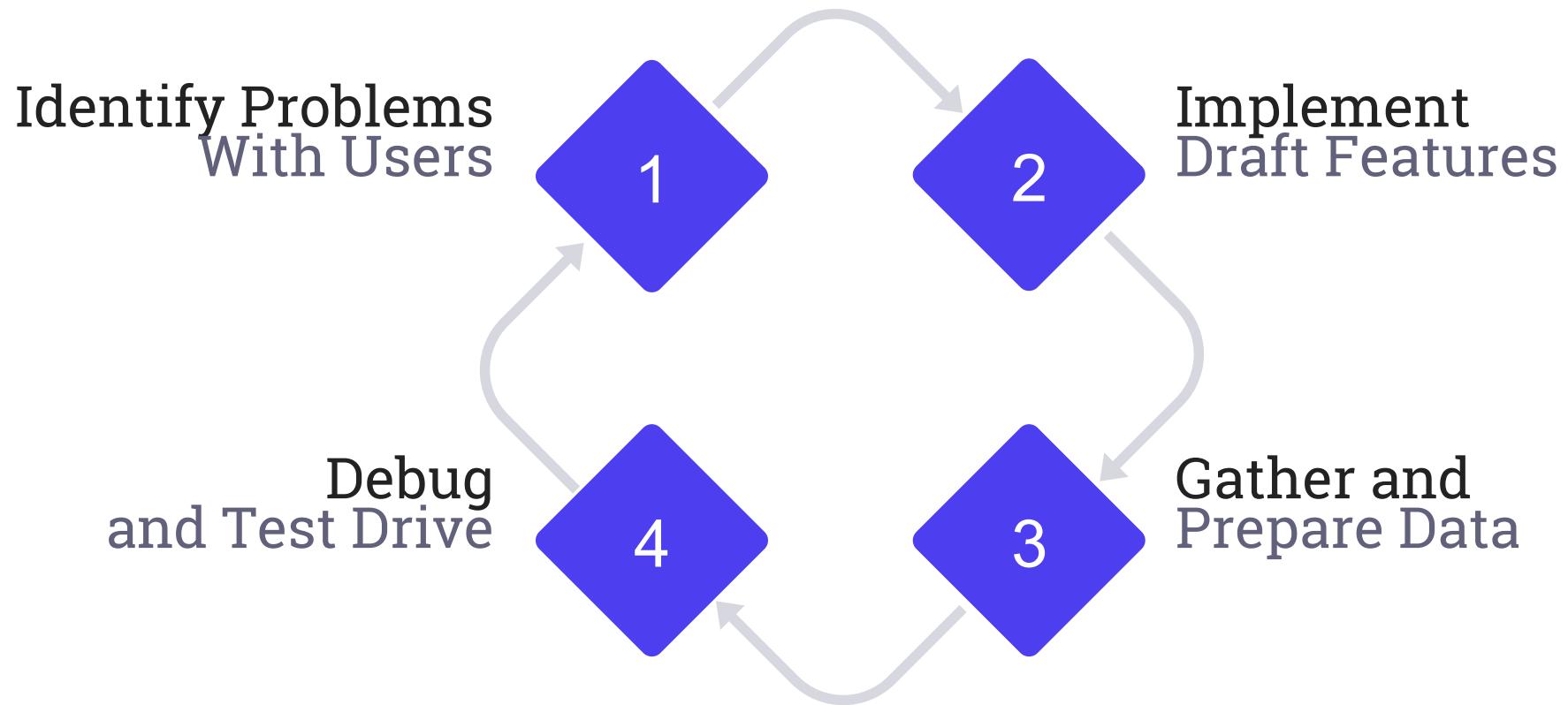


Try Everything  
Fail Fast



User  
Experience

# Development Cycle



# Track MEA



A web application that tracks events that happen throughout a MEA's lifetime.

#Mass Neuron Manufacturing, etc  
Input from Ben, Isabella

# USE CASES

- > **Neurons from an entire batch of MEAs just died.**  
Are there any common events in these MEA's history that may explain this anomaly? A common media preparation? How can I optimize for MEA longevity?
- > **We just got a new shipment of MEAs.**  
Before mixing them with the old MEAs, we need to keep track of when and where we got them.
- > **I want the newest MEAs for this experiment.**  
Which MEAs have only been in operation for less than a month? Likewise, get arbitrary MEA groupings with the conditions that you need.
- > **I collected a bunch of electrical data.**  
Correlate MEA response properties (detection features, signal-to-noise) with factors within this MEA's setup (OR, media) and previous lifespan.

# Workflow



Generate QR  
code with Python

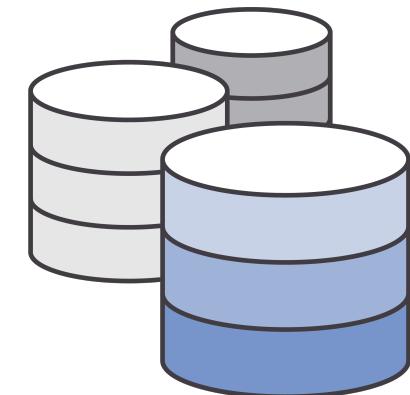


Life Events

|    | A            | B       | C            | D               | E               |
|----|--------------|---------|--------------|-----------------|-----------------|
| 1  | Coating      | Seeding | Density      | Media           | Gene Delivery   |
| 2  | PDL 50ug/ml  | NSC     | 5E5 cells/ml | BPB+            | Lenti           |
| 3  | PDL 100ug/ml | OSC     | 2E6 cells/ml | BPB-            | AAV             |
| 4  | PDL 5mg/ml   | ODR     |              | Differentiation | Lipo2000        |
| 5  | PEI 100ug/ml | PNC     |              |                 | Lipo3000        |
| 6  | PEI 5mg/ml   | H3A     |              |                 | Neurofect       |
| 7  | Laminin      | HEK     |              |                 | Neuroporter     |
| 8  | None         | HET     |              |                 | FuGene8         |
| 9  |              |         |              |                 | FuGeneHD        |
| 10 |              |         |              |                 | Electroporation |
| 11 |              |         |              |                 | CaPO4           |

MEA Tracking

|    | A      | B            | C                | D       | E       |
|----|--------|--------------|------------------|---------|---------|
| 1  | MEA id | Lot id       | Timestamp        | Coating | Seeding |
| 2  |        | PNC-20180903 | 09/03/18 01:06PM |         | PNC     |
| 3  |        | PNC-20180903 | 09/03/18 01:07PM |         | PNC     |
| 4  |        | PNC-20180903 | 09/03/18 01:09PM |         | PNC     |
| 5  |        | PNC-20180903 | 09/03/18 01:10PM |         | PNC     |
| 6  |        | PNC-20180903 | 09/03/18 01:11PM |         | PNC     |
| 7  |        | PNC-20180903 | 09/03/18 01:11PM |         | PNC     |
| 8  |        | PNC-20180903 | 09/03/18 01:12PM |         | PNC     |
| 9  |        | PNC-20180903 | 09/03/18 01:12PM |         | PNC     |
| 10 |        | PNC-20180903 | 09/03/18 01:13PM |         | PNC     |
| 11 |        | PNC-20180903 | 09/03/18 12:10PM |         |         |



MEAs  
Cultures  
Events

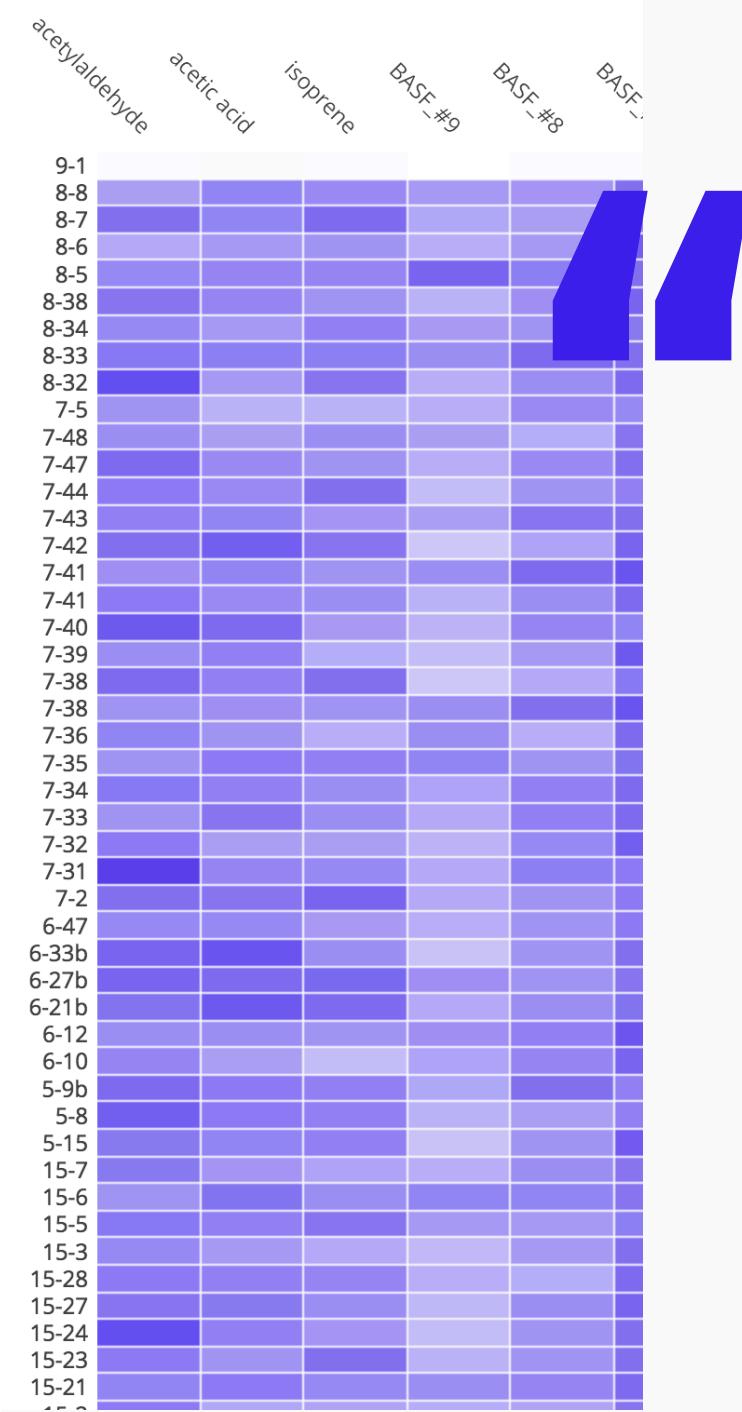
Form results stored  
into database



# TrackMEA DEMO



# Odor Receptor Database (ORDB)



A user-friendly, up-to-date, searchable knowledge engine of Koniku's OR and Compound data collected from screening and dose response experiments.

#ORDB  
Input from Mark, Christian, Amit



# USE CASES

---

- > Give me all the experiments done on this OR / this Compound.  

ORDB can perform key-word searches on OR and Compound names to quickly find your target experiments.
  - > What is the Limit of Detection of OR19-20 for Acetophenone?  

Import raw data from Dose Response experiments, generating heatmaps and dose response curves, immediately showing the limit of detection.
  - > Which of these ORs is the most sensitive to TNT?  

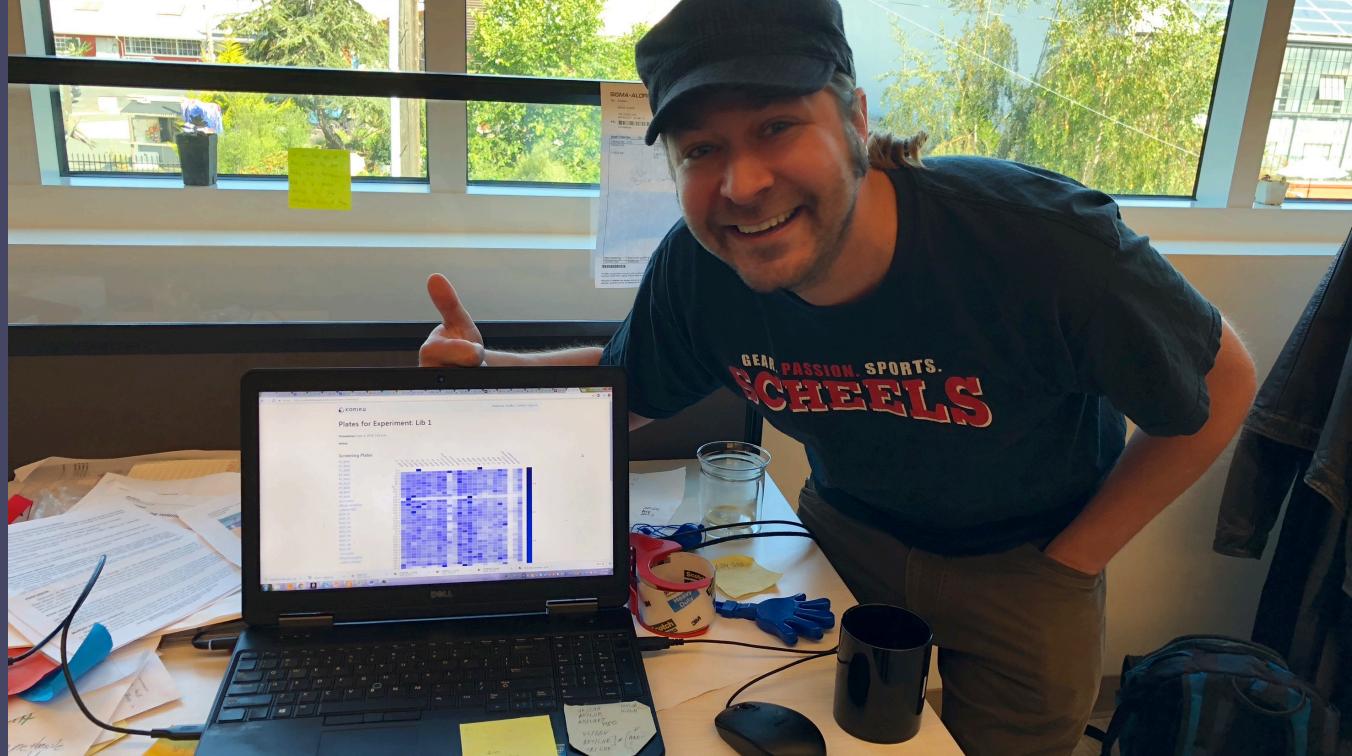
Import raw data from Screening Experiments, generating heatmaps between ORs and Compounds. It's never been easier to identify hits immediately!
  - > Which compound produces the closest collective OR response to Norcocaine?  

OR-compound matrices allow you to easily cluster similar compounds / similar ORs, key for Perceptome projects.
- Featuring OdorDB:
- > Which VOCs are associated with sweat?  

OdorDB allows you to search for VOCs that are associated with odors that matter for your life and your industry. These VOCs may be detected by Koniku's technology.

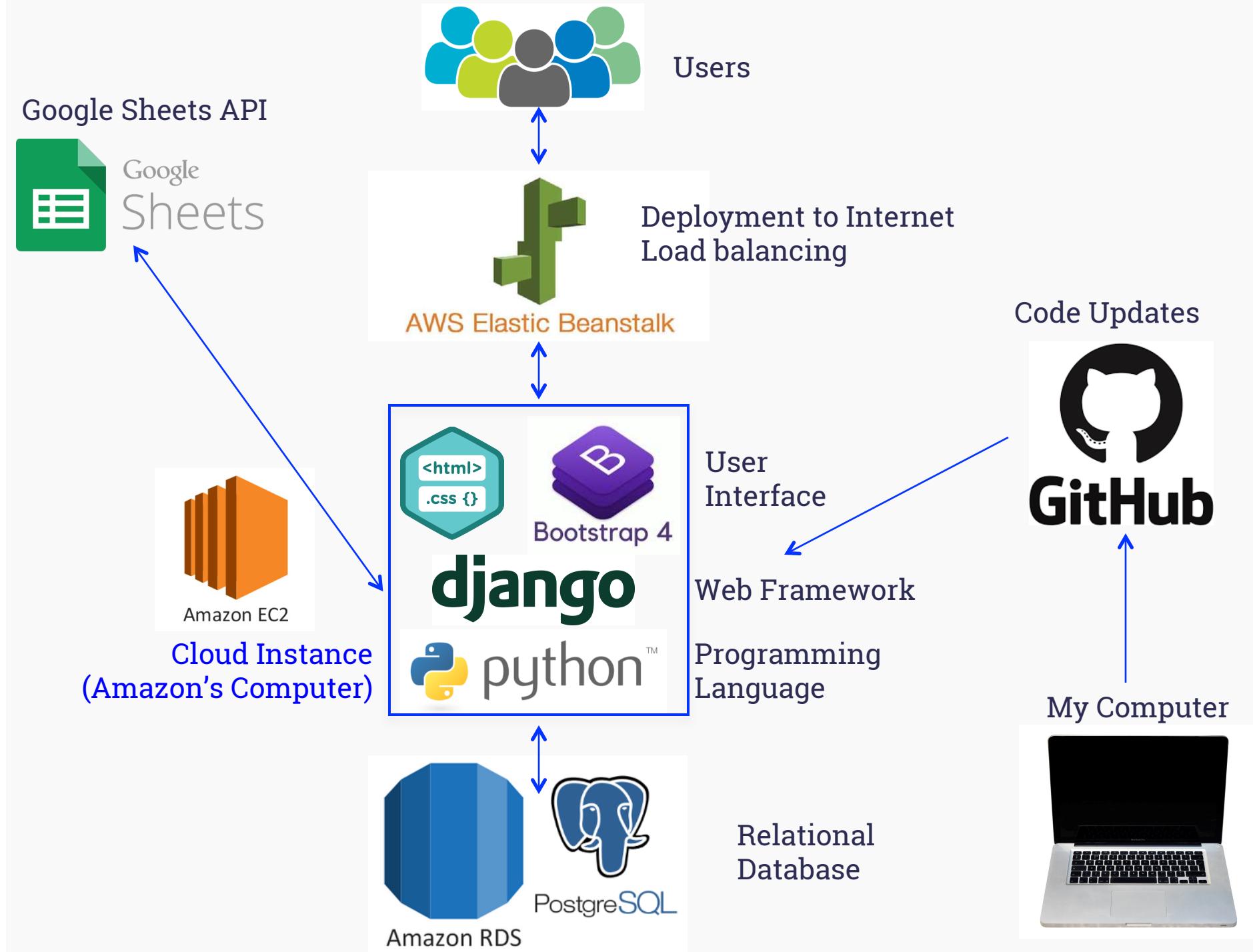


# ORDB DEMO

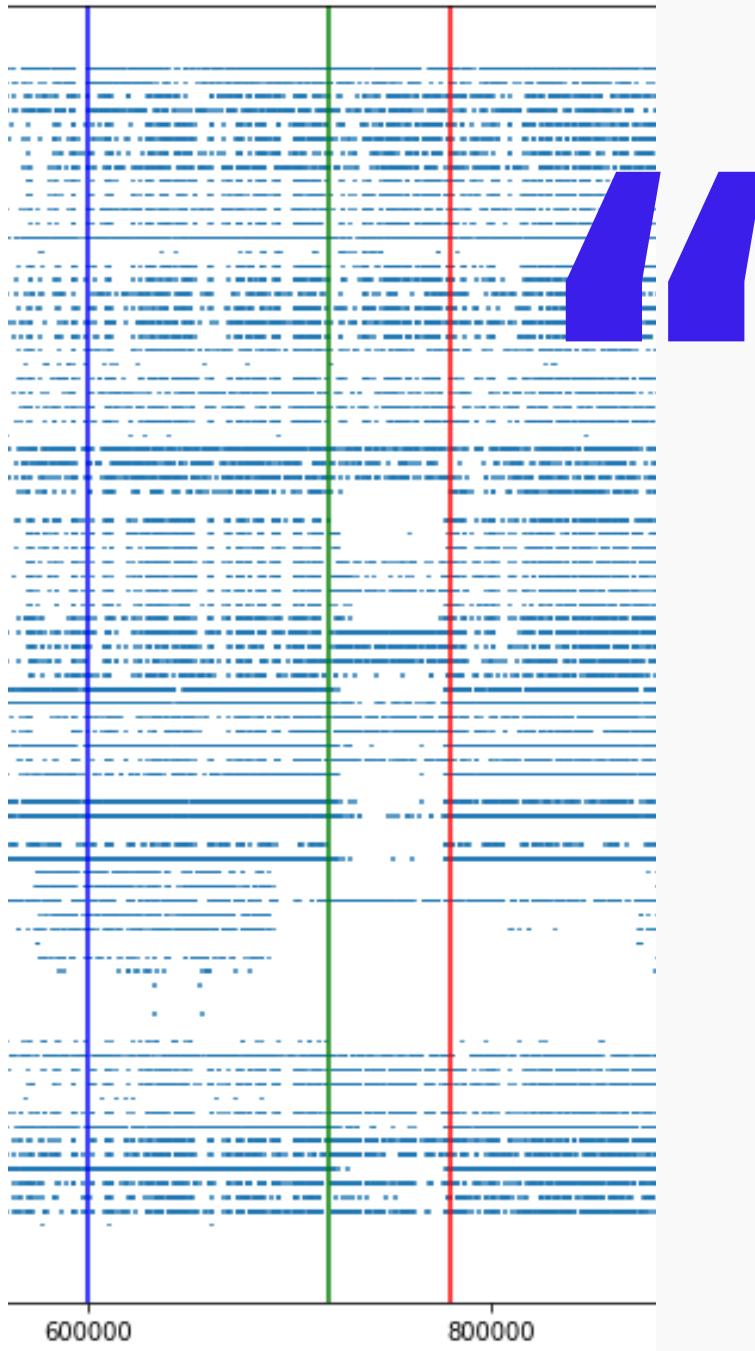




# Web App Stack



# MEA Data Pipeline

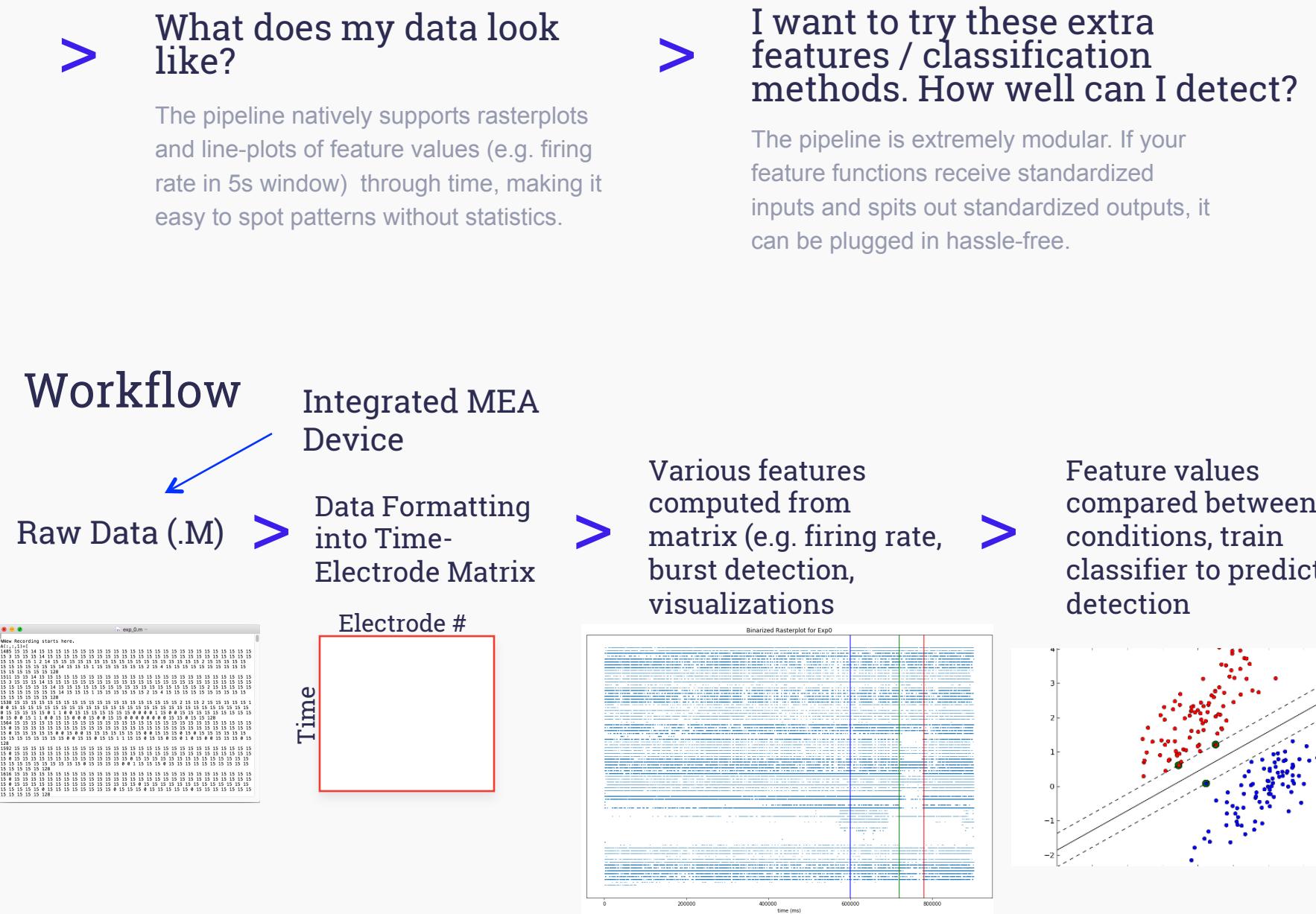


A modular data ingestion, processing, visualization, and analysis pipeline for MEA data. End goal is to build an algorithm to detect features within MEA data that will indicate a detection.

#NERO

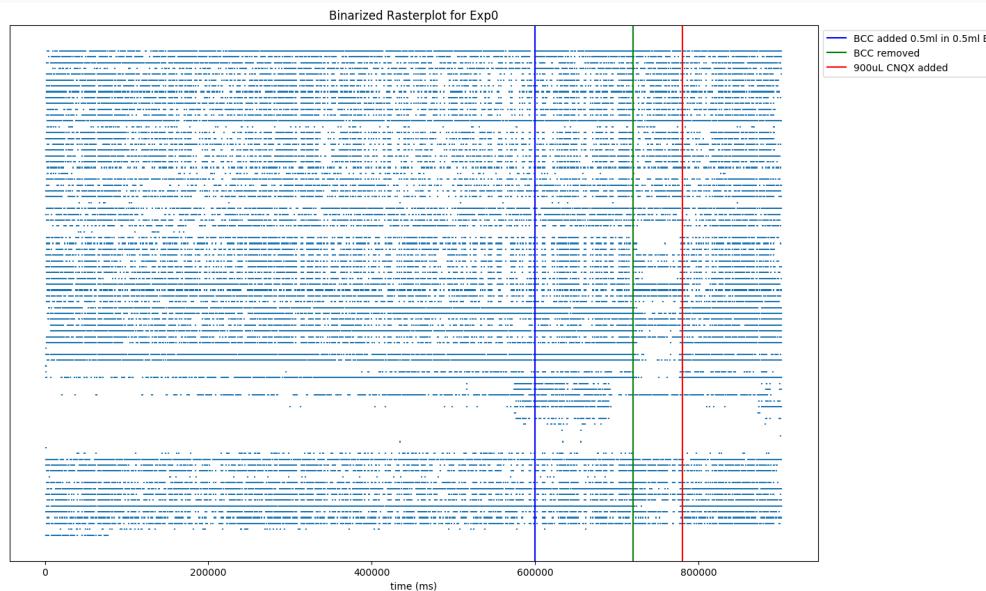
Data input from Winston, Anisha  
Code input from Marius

# USE CASES & WORK FLOW

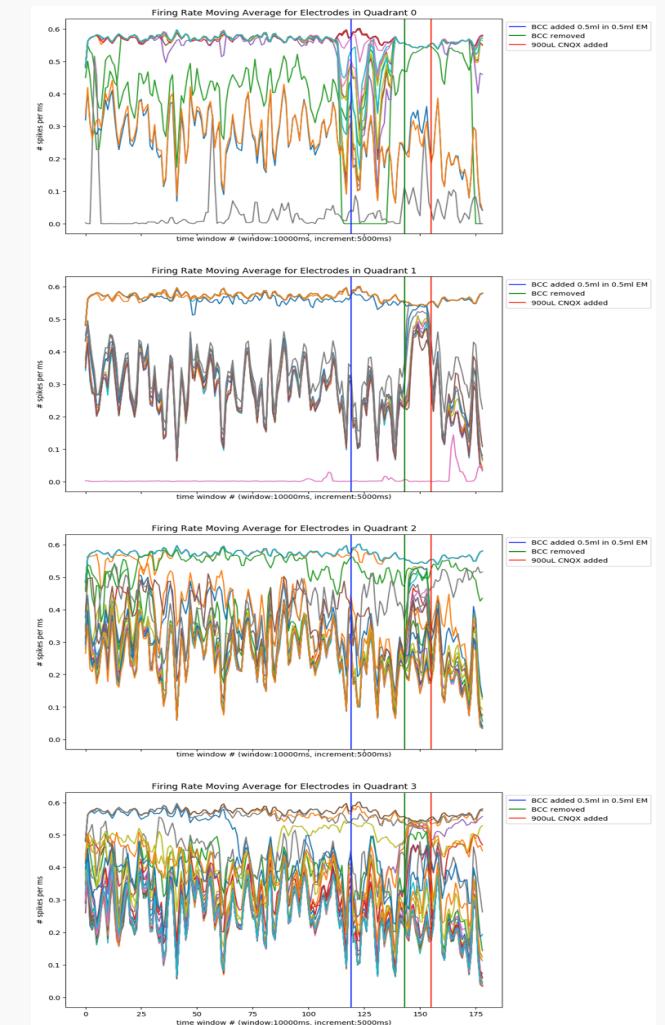


# Gallery

Rasterplot with 3 events

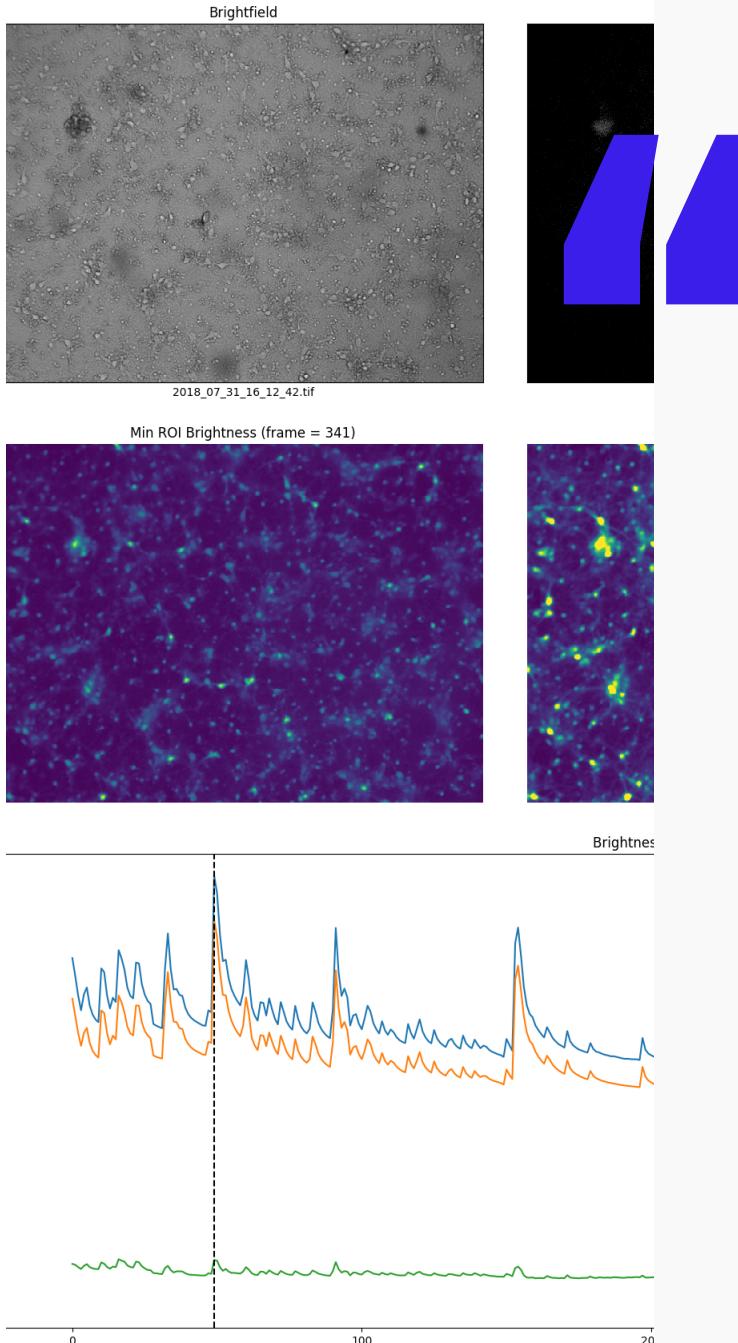


Moving average of Firing Rate for electrodes in each quadrant



# Calcium Imaging Data Pipeline

| Lot  | Plate | Well | Days In Vitro | Days w/ Virus | OR    |
|------|-------|------|---------------|---------------|-------|
| NC46 | 11    | F6   | 20            | 18            | Naive |



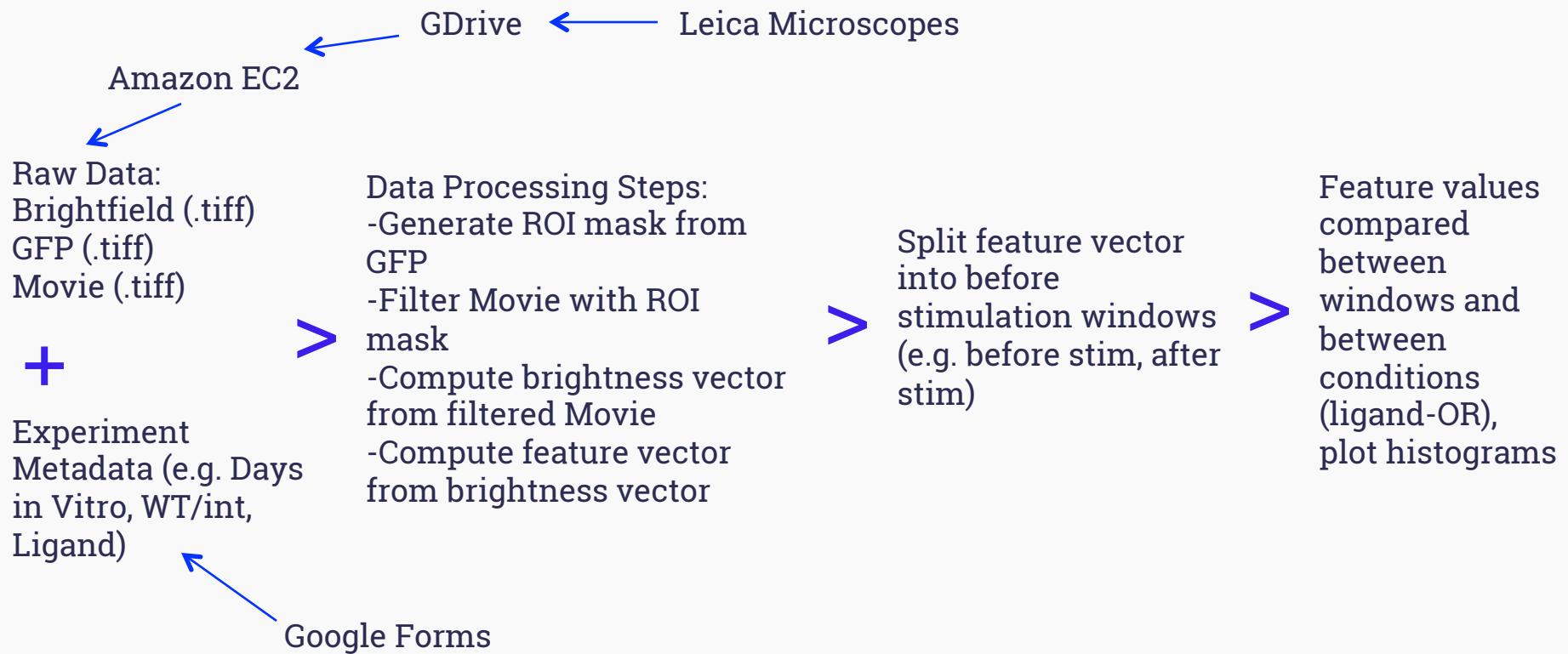
A common data ingestion, processing, visualization, and analysis pipeline for all calcium imaging experiments. End goal is to build an algorithm to detect features within calcium imaging data that will indicate a detection.

#ReveLOR #InsectOR #IODE  
Data input from everyone in above cells  
Form input from Anisha, Vivek  
Code input, infrastructure setup from Marius



# USE CASES & WORK FLOW

## Workflow

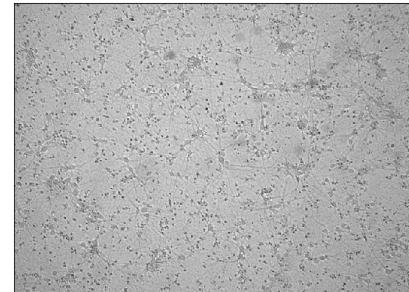




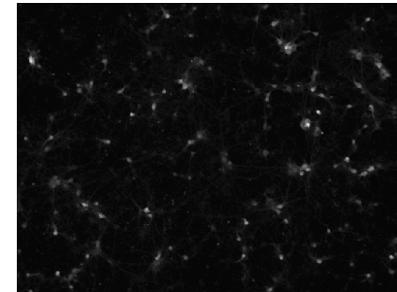
## Comprehensive Summary Plot for a ReveLOR Experiment

| Plate | Well | 1 Ligand | Lot  | Days w/ Virus | OR  | 1 Stim Time | WT/int | CNG  | 1 Conc (uM) | Leica | Magnification | Protein Tag | Days In Vitro |
|-------|------|----------|------|---------------|-----|-------------|--------|------|-------------|-------|---------------|-------------|---------------|
| n12   | F3   | RDX      | NC46 | 18            | 185 | 120         | N/A    | None | 200         | 2     | 10X           | GFP         | 20            |

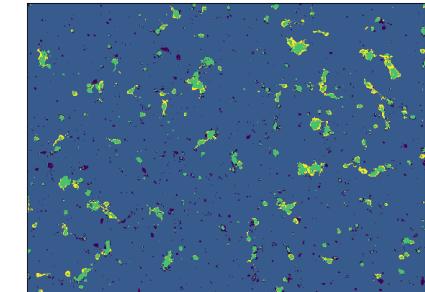
Brightfield



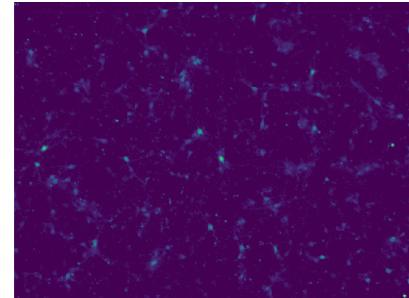
GFP



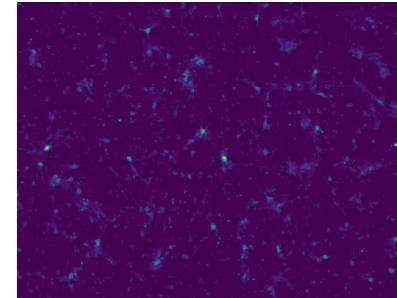
ROI



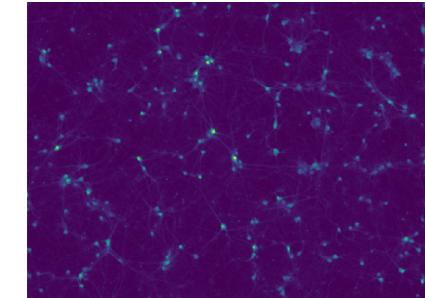
Min ROI Brightness (frame = 197)



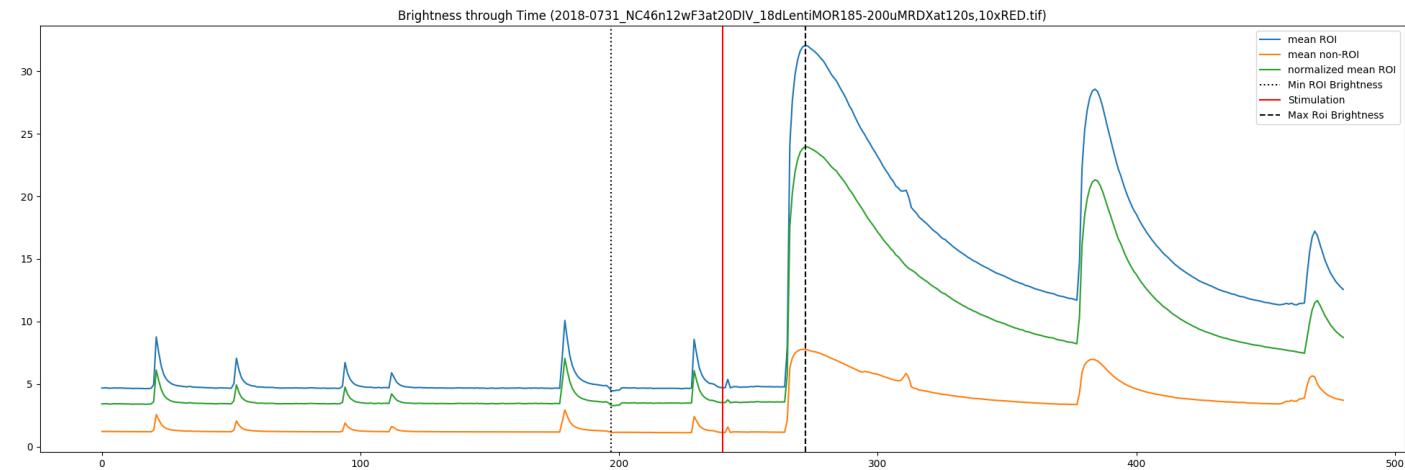
Stimulation (frame = 240)



Max Roi Brightness (frame = 272)



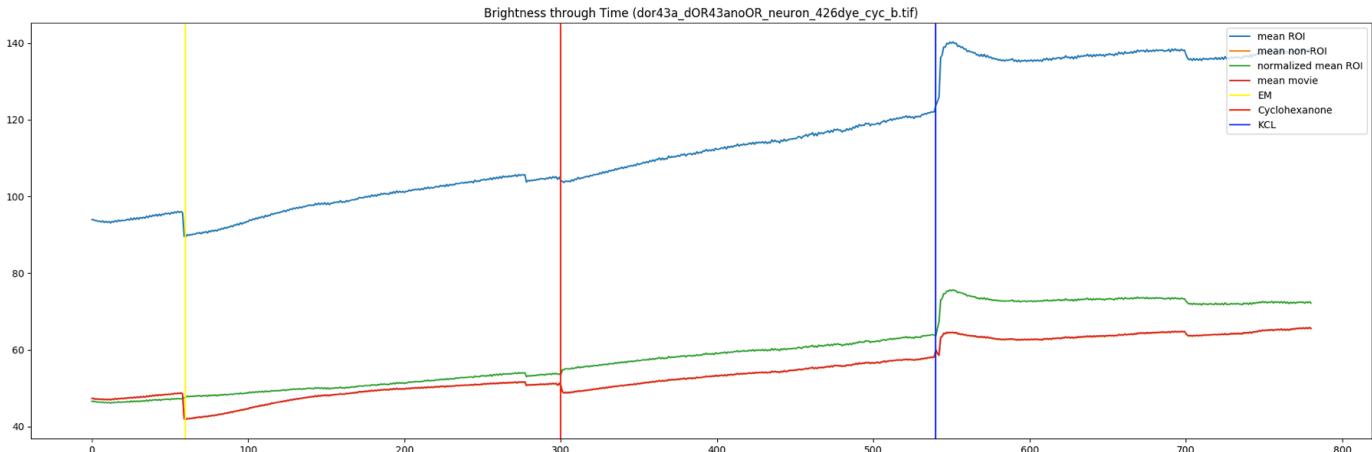
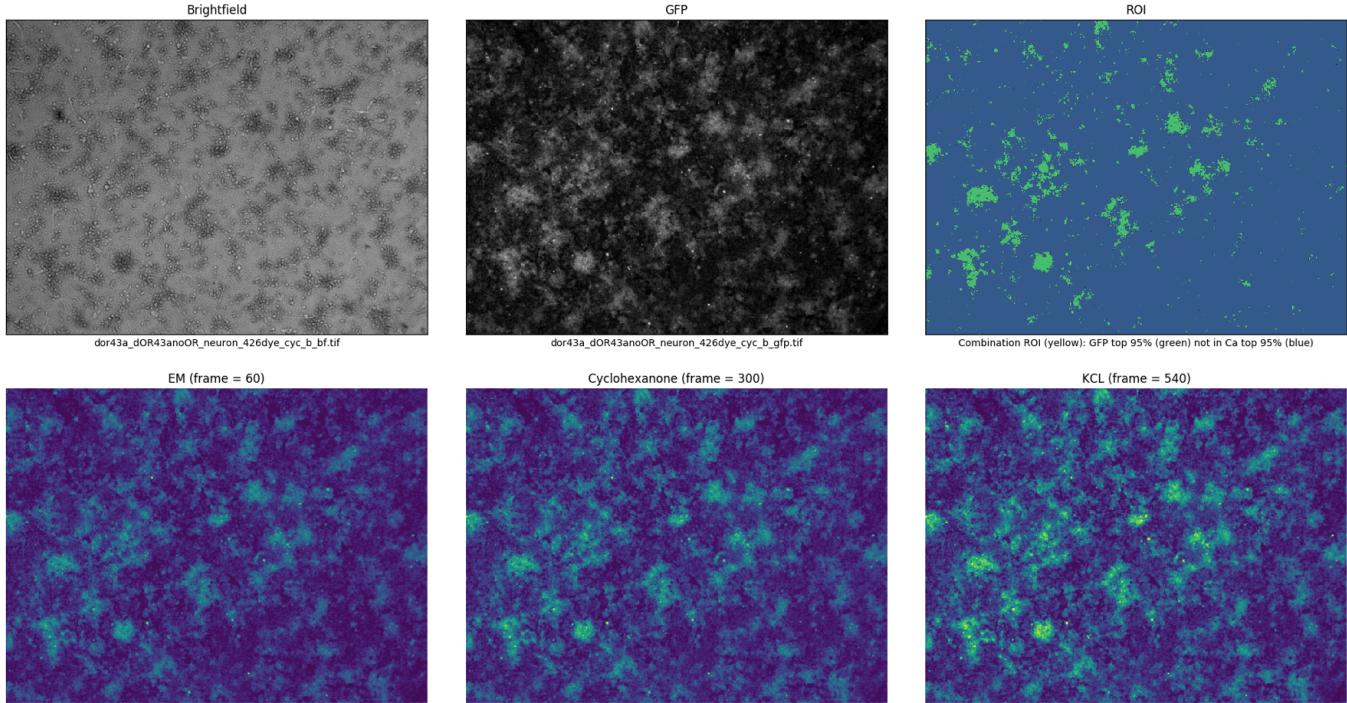
# Gallery





## Comprehensive Summary Plot for a InsectOR Experiment

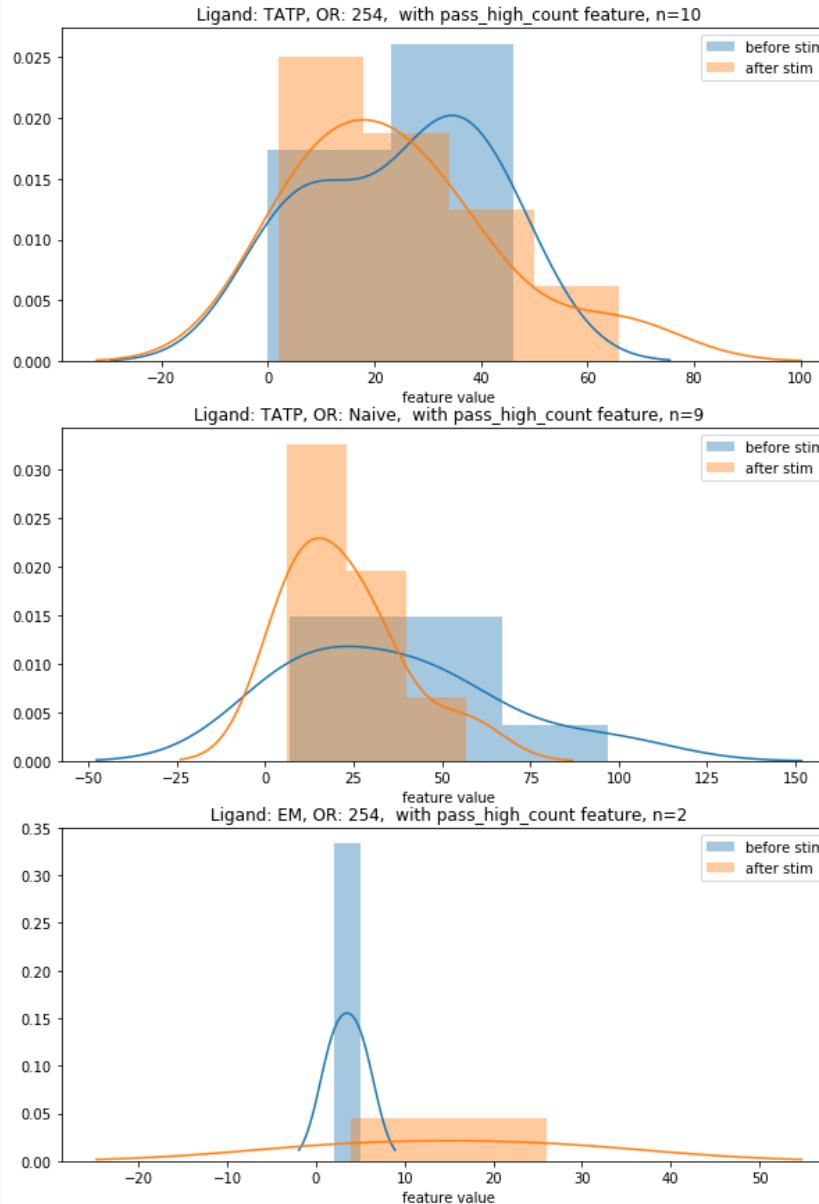
| Protein Tag | WT/int | Magnification | Lot | Plate | Leica | Days In Vitro | Ligand 1 Cconcentrations(uM) | Well | CNG  | OR    | Days w/ Virus | StimTimes(s) |
|-------------|--------|---------------|-----|-------|-------|---------------|------------------------------|------|------|-------|---------------|--------------|
| GFP         | wt     | 10X           | 1   | 4     | 1     | 19            | EM                           | c10  | None | Naive | 8             | 30, 150, 270 |



# Gallery

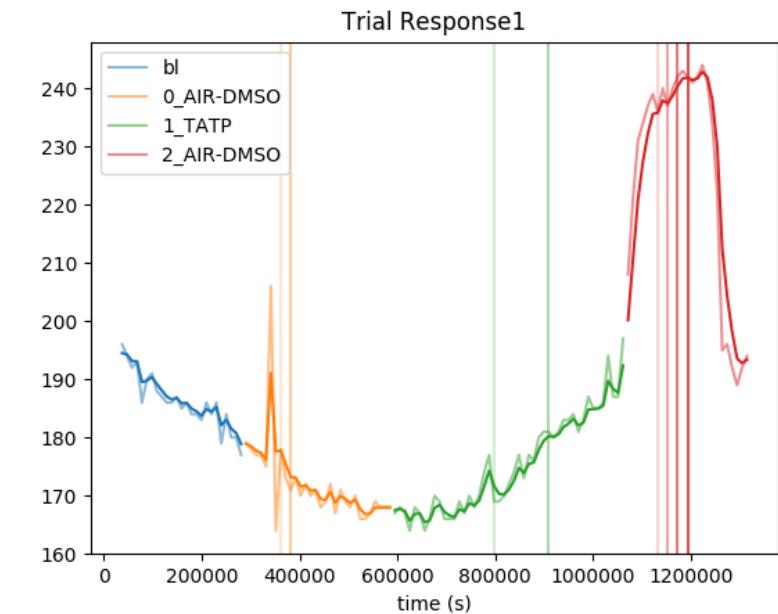
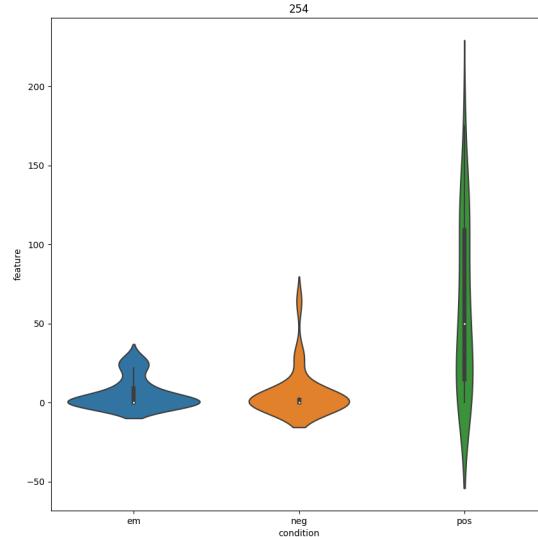
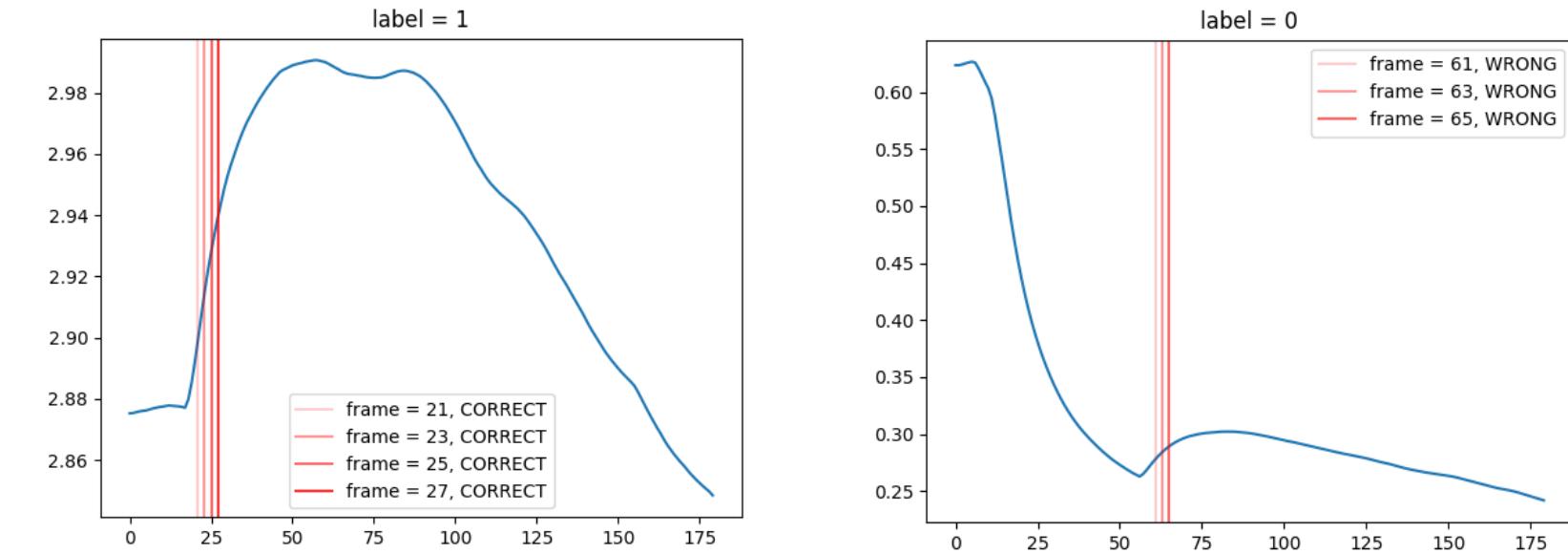
# Gallery

Histograms of feature values before/after stimulation,  
For experiment and control conditions in ReveLOR



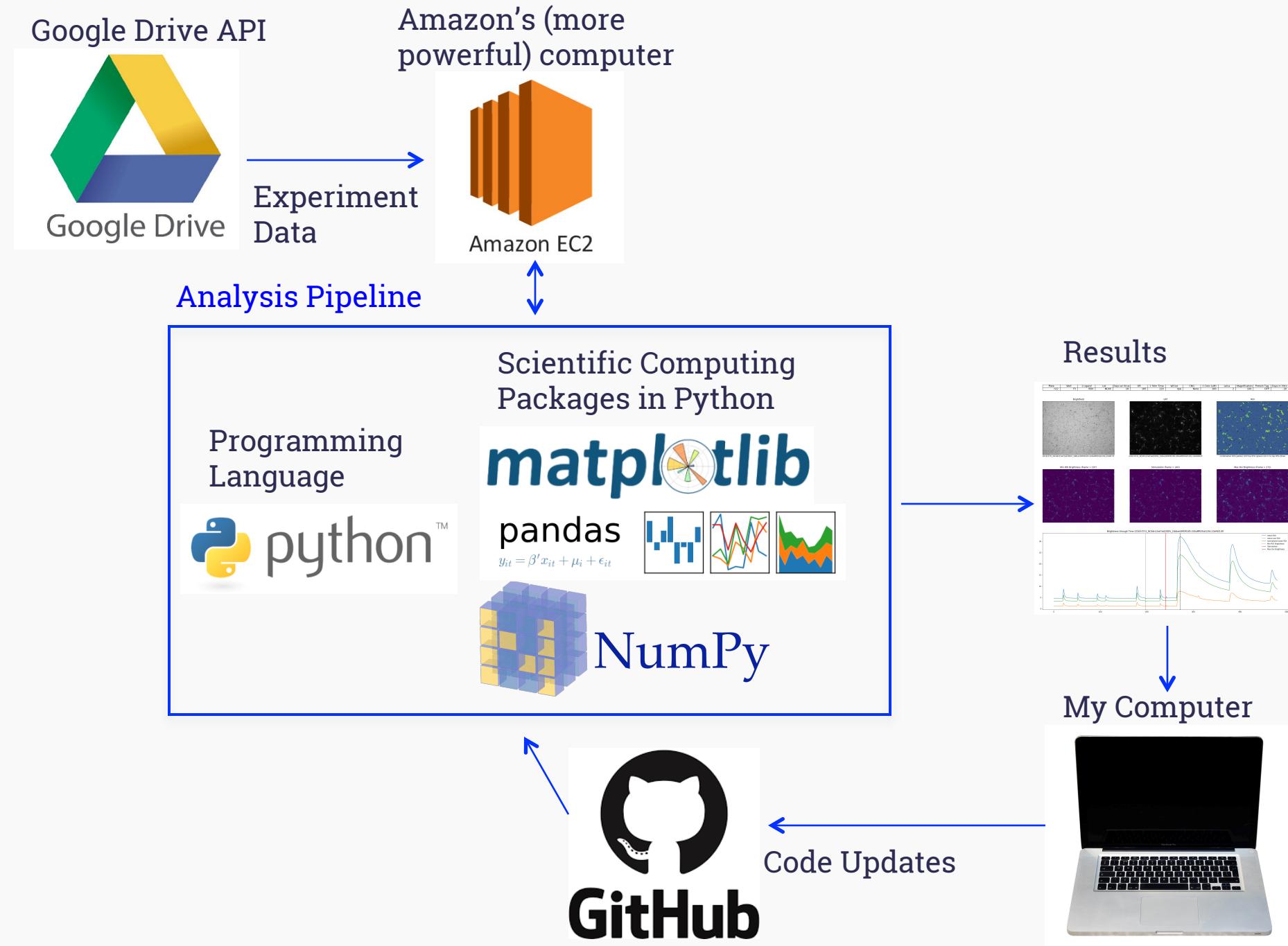
## Histograms of feature values before/after stimulation, For experiment and control conditions.

# Gallery





# Data Pipeline Stack



# That's a lot of Code!

## ORDB Private

● Python Updated an hour ago



## trackit Private

MEA tracking web app based on Django, with Google Sheets backend

● Python Updated 5 hours ago



## iode-analysis Private

calcium imaging analysis of IODE data

● Python Updated a day ago



## revelor-analysis Private

● Python Updated 17 days ago



## mea\_classification Private

Data analysis of MEA data for the purpose of classification

● Python Updated on Aug 1



## OdorDB Private

● Python Updated on Jul 23





# Misc. Work

- > Filed a patent?!  

Filed a patent with Osh and Marius on the Perceptome Generation Engine – a paradigm to systematically collect human behavioral data when exposed to odorants and correlate that with the patterns produced by the Koniku Kore.
- > Made a software demo video for Henkel.  

Worked with Osh and used features of both OdorDB and ORDB to demonstrate a potential software platform for the Koniku Kore.
- > Interviewed people for a full time job?!  

Together with Marius, we interviewed ~10 people for a Data Scientist full-time job. Only at a dope startup do interns get to do this!
- > Set up an EC2 Server (running Windows, ugh.) for MesaLabs software.  

Now our fridges are going to be connected to the cloud, wow. Along with Winston, we are moving closer to Koniku's smart-lab-on-the-cloud dream.



# I learned a lot.

- > **Most of the things in the tech stacks I've mentioned.** >
  - I have never worked with things like Django, AWS, Google Drive APIs before, and picked them up from scratch. As my projects required more functionality, I learned more methods to solve these problems.
- > **How to communicate with users.** >
  - Repeated the Development Cycle many times with the people who will be using my applications. Learned how to solicit and apply feedback.
- > **Lots of startup knowledge.**
  - I knew terminology about funding rounds, VCs, etc, but I learned so much more about startups when working in one. Team dynamics, company market strategy, etc... Thanks for being so transparent!
- > **Working in a team.**
  - Thanks to the Cell System, I got to work in many different teams trying to solve really interesting problems. I learned so much from everyone. Special thanks to Marius for being a great mentor!



THANK YOU  
MERCI  
谢谢