



# Machine Learning for CI

Kyra Wulffert

[kwulffert@yahoo.com](mailto:kwulffert@yahoo.com)

Matthew Treinish

[mtreinish@kortar.org](mailto:mtreinish@kortar.org)

Andrea Frittoli

[andrea.frittoli@gmail.com](mailto:andrea.frittoli@gmail.com)

August 29, 2018

[https://github.com/afrittoli/ciml\\_talk](https://github.com/afrittoli/ciml_talk)



# CI at Scale

- ▶ Continuous Integration
- ▶ Continuous Log Data
- ▶ Lot of data, little time
- ▶ Triaging failures?
- ▶ AI to the rescue!

## The OpenStack use case

- ▶ Integration testing in a VM
- ▶ System logs, application logs
- ▶ TBD pic of dstat data
- ▶ Dstat data
- ▶ Not only OpenStack

## Building the dataset

- ▶ TBD Automation in collecting data
- ▶ Object storage / Openwhisk

## Infrastructure for training

- ▶ Helm deployment
- ▶ Ffdl <- s3 input
- ▶ Experiment / dataset in Cloudant / via OpenWhisk API

## Infrastructure for prediction

- ▶ MQTT
- ▶ API / OpenWhisk

## Data stability

- ▶ Dataset selection
- ▶ Reproducibility

Thank you!  
Questions?