# Kubeflow End-to-End: GitHub Issue Summarization

Amy Unruh
Michelle Casbon
Kubecon North America
Seattle
December 11, 2018
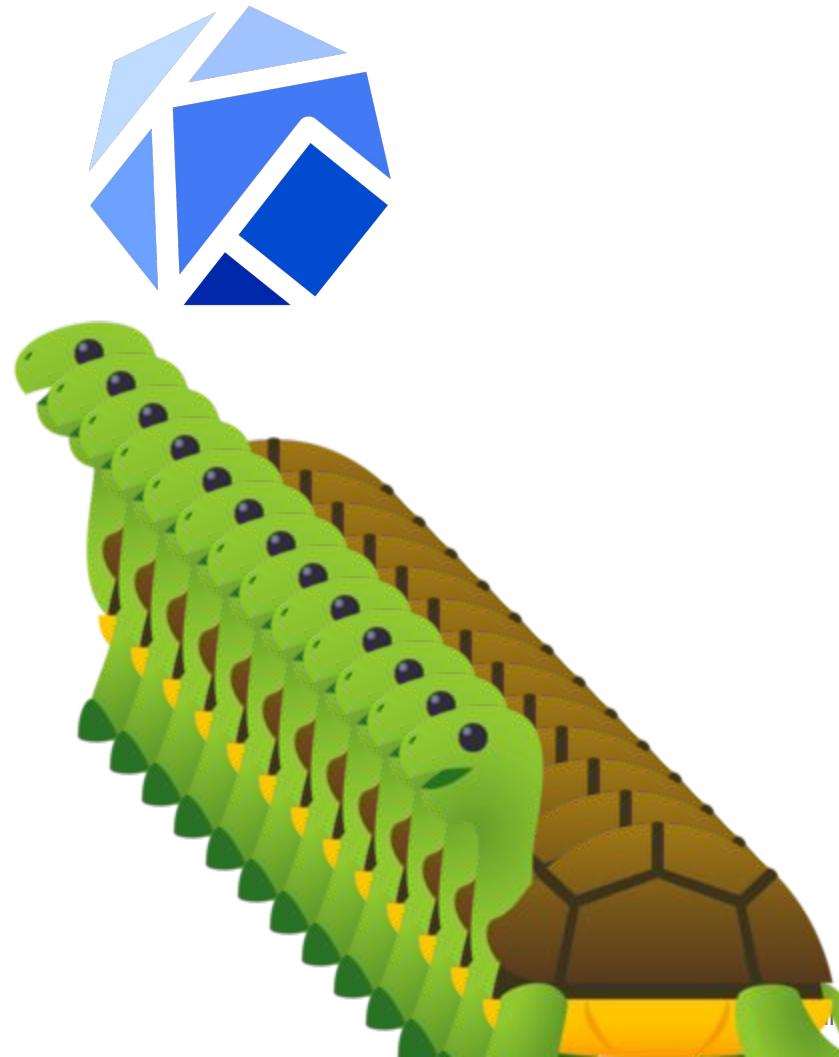
Google Cloud

Google Cloud
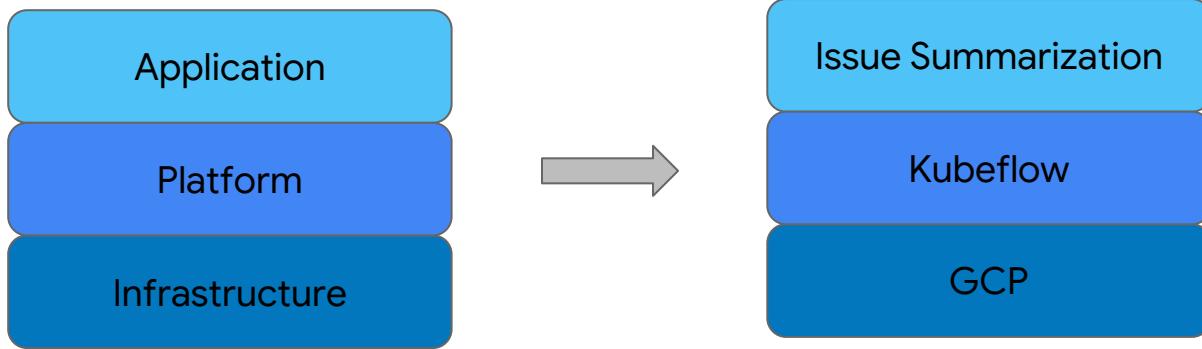
# Kubeflow is a curated set of compatible tools and artifacts that lays a foundation for running production ML apps

Enables consistency across deployments by providing Kubernetes object templates that bring together disparate components

Google Cloud

@amygdala @texasmichelle

# Contributors

- Kubeflow is open
  - Open community
  - Open design
  - Open source
  - Open to ideas
- Get involved
  - github.com/kubeflow
  - kubeflow.slack.com
  - @kubeflow
  - kubeflow-discuss@googlegroups.com
  - Community call Tuesdays alternating 8:30am and 5:30pm Pacific
  - **Kubeflow Contributor Summit**
    - Q1 2019

Google Cloud

`https://github.com/kubeflow/kubeflow`

@amygdala @texasmichelle

# Agenda   g.co/codelabs/kubecon18

**1** — Set up the environment

**2** — Create a Kubeflow cluster

**3** — Run a pipeline from the Kubeflow Pipelines dashboard

**4** — Run a pipeline from a Jupyter notebook

**5** — Clean up

Google Cloud
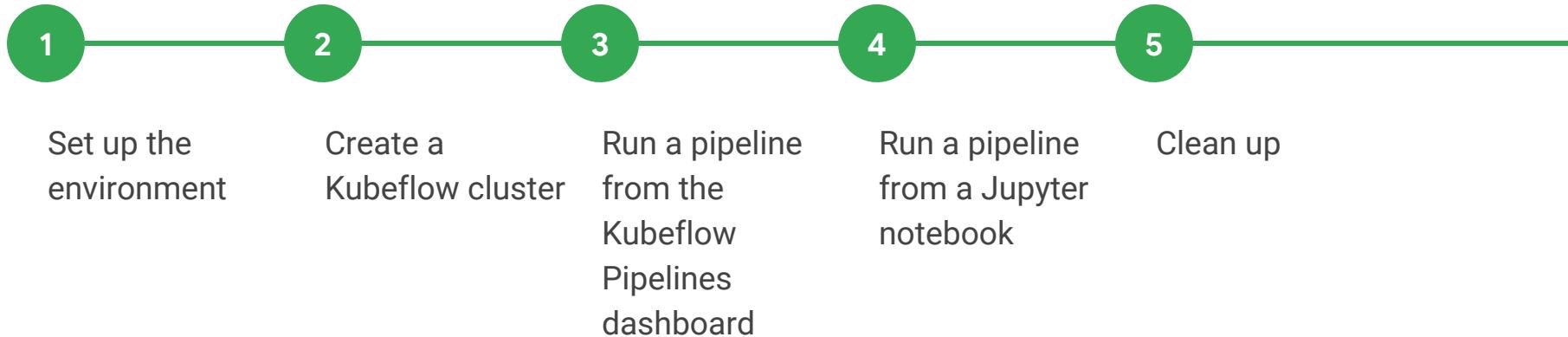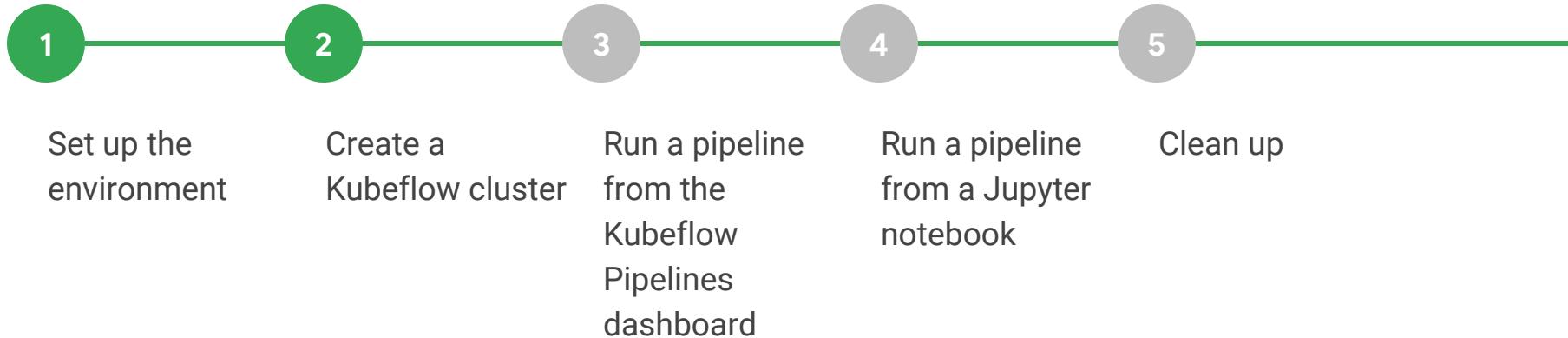
# Agenda    g.co/codelabs/kubecon18

Zones: us-west1-b, us-east1-c

**1** — Set up the environment

**2** — Create a Kubeflow cluster

**3** — Run a pipeline from the Kubeflow Pipelines dashboard

**4** — Run a pipeline from a Jupyter notebook

**5** — Clean up

Google Cloud

PROBLEM

# Moving from local to production



SOLUTION

## Portability

Package infrastructure components together

Google Cloud

Credit: Jörg Wagner and Stefan Prehn

@amygdala @texasmichelle

PROBLEM
# Complexity

Google Cloud

@amygdala @texasmichelle

# Perception

ML Code

| | | | | | |
|---|---|---|---|---|---|
| Configuration | Data Collection | Data Verification | Machine Resource Management | | Monitoring |
| | Feature Extraction | Analysis Tools | Serving Infrastructure | | |
| | | Process Management Tools | | | |

Credit: Hidden Technical Debt of Machine Learning Systems, D. Sculley, et al.

Google Cloud

@amygdala @texasmichelle

# Reality

GitHub IS
Kubeflow
GCP

Configuration

Data Collection

Data Verification

ML Code

Machine Resource Management

Analysis Tools

Feature Extraction

Process Management Tools

Serving Infrastructure

Monitoring

Credit: Hidden Technical Debt of Machine Learning Systems, D. Sculley, et al.

Google Cloud

@amygdala @texasmichelle

# Data

- Data Ingestion
- Data Exploration
- Data Transformation
- Data Validation
- Data Analysis
- Training Data Segmentation

# Featurization

- Feature Extraction

# Training

- Model Building
- Model Validation
- Model Versioning
- Model Auditing
- Distributed Training
- Continuous Training

# Application

- Serving Infrastructure
- Business Logic
- UI
- Load Balancing

# Platform

GitHub IS
Kubeflow
GCP

- Configuration
- Process Management
- Resource Management
- Monitoring
- Logging
- Continuous Delivery
- Authentication/ Authorization

Google Cloud

@amygdala @texasmichelle

PROBLEM
# Complexity

SOLUTION
## Composability

Logical groupings

Reusable components

Google Cloud

PROBLEM

# Maintainability

- Error resolution, recovery, & prevention
- Speed of iteration
- Versioning

SOLUTION

## Composability

Shorten the development lifecycle

Automation

Google Cloud

PROBLEM

# Capacity Planning
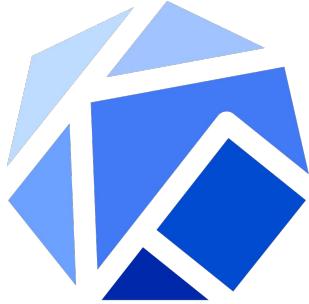
- Usage patterns
- Demand spikes
- Efficient resource usage

SOLUTION

## Scalability

Kubernetes

Autoprovisioning

Google Cloud

@amygdala @texasmichelle

Make it easy for everyone to develop, deploy, and manage portable, scalable ML everywhere

Google Cloud

# Kubeflow

**Portability**

Entire stack

**Scalability**

Native to k8s

Reduce variability between services & environments

**Composability**
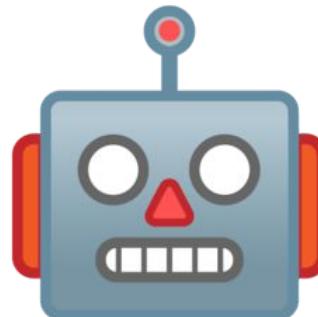
Single, unified tool for common processes

**Full product lifecycle**

**Support specialized hardware, like GPUs & TPUs**

Reduce costs

Improve model performance

Google Cloud

@amygdala @texasmichelle

# Kubeflow

**Who**

Data scientists

ML researchers

Software engineers

Product managers

**What**

Portable ML products on k8s

v0.3.4 release

**Why**

Because building a platform is too big of a problem to tackle alone

`https://github.com/kubeflow/kubeflow`

Google Cloud

# Kubeflow

**Kubernetes-native platform for ML**

Run wherever k8s runs

Use k8s to manage ML tasks

CRDs for distributed training

**Adopt k8s patterns**

Microservices

Manage infra declaratively

**Package infrastructure components together**

Ksonnet

Move between local -> dev -> test -> prod -> onprem

**Support multiple ML frameworks**
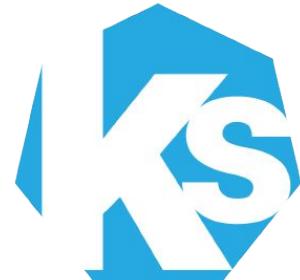
TensorFlow

Pytorch

Scikit

Xgboost

Et al.

Google Cloud

@amygdala @texasmichelle

# Agenda    g.co/codelabs/kubecon18

**1** Set up the environment

**2** Create a Kubeflow cluster

**3** Run a pipeline from the Kubeflow Pipelines dashboard

**4** Run a pipeline from a Jupyter notebook

**5** Clean up

Google Cloud

# What's Inside v0.3?



@amygdala @texasmichelle
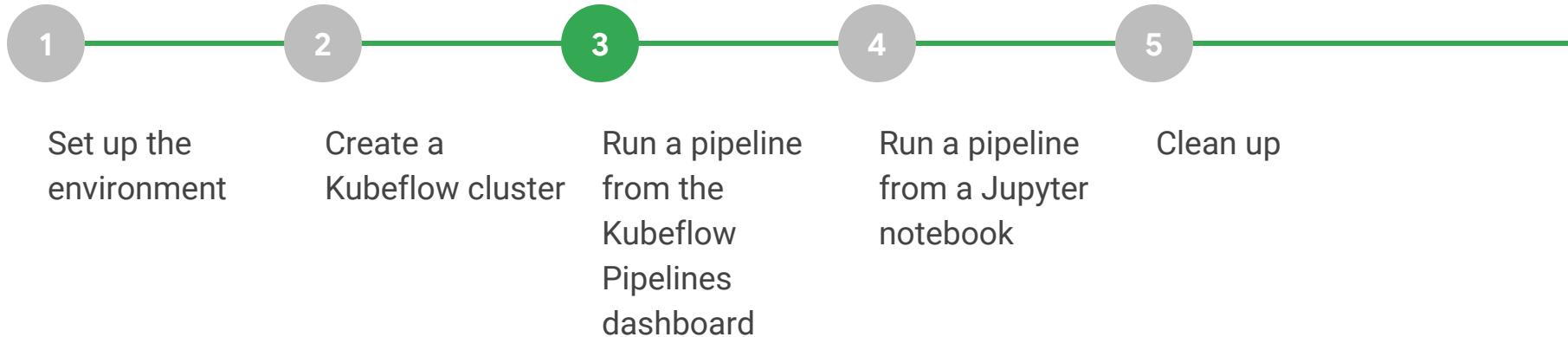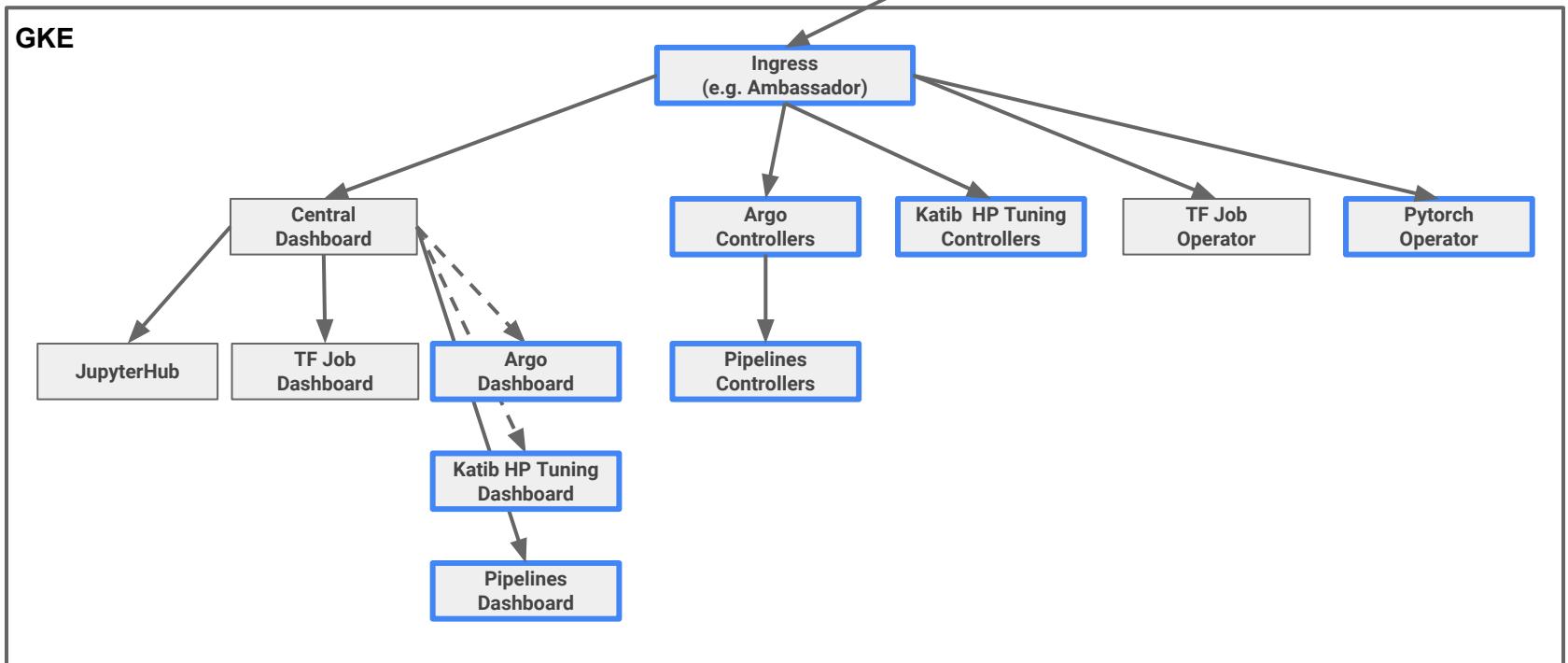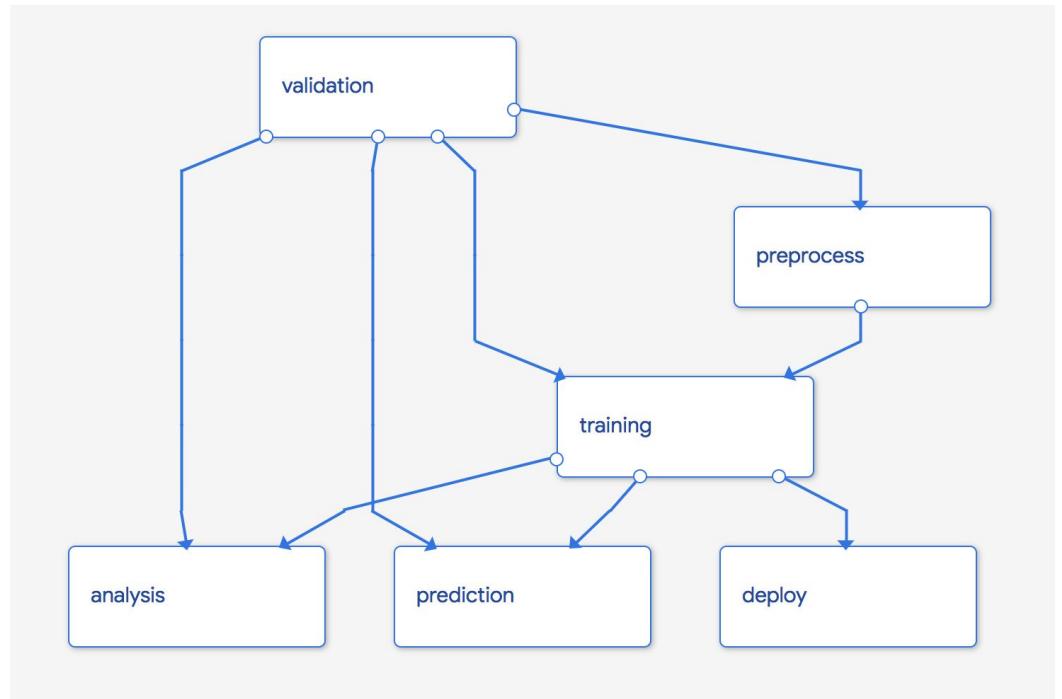
# What's new in v0.3?

- Deploy
  - Click-to-deploy
  - CLI (`kfctl`)
- Develop
  - Argo
  - Pytorch operator
  - Hyperparameter tuning StudyJob CRD
  - Kubeflow Pipelines

Google Cloud
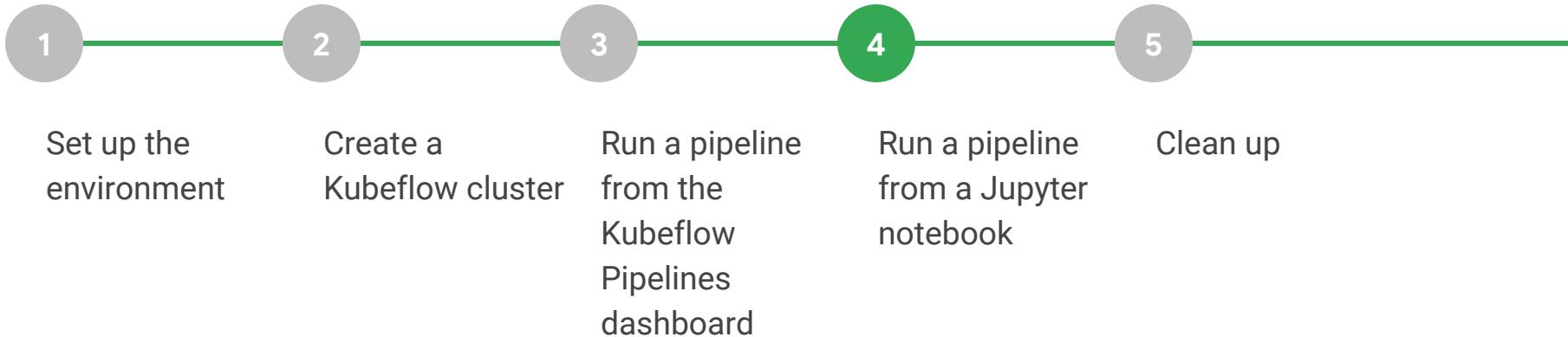
# Pipelines

- End-to-end ML workflows
- Orchestration
- Service integration
- Components & sharing
- Job tracking, experimentation, monitoring
- Notebook integration



Google Cloud

@amygdala @texasmichelle

# Agenda   g.co/codelabs/kubecon18

**1** — Set up the environment

**2** — Create a Kubeflow cluster

**3** — Run a pipeline from the Kubeflow Pipelines dashboard

**4** — Run a pipeline from a Jupyter notebook

**5** — Clean up

Google Cloud

# Agenda   g.co/codelabs/kubecon18

**1** — Set up the environment

**2** — Create a Kubeflow cluster

**3** — Run a pipeline from the Kubeflow Pipelines dashboard

**4** — Run a pipeline from a Jupyter notebook

**5** — Clean up

Google Cloud
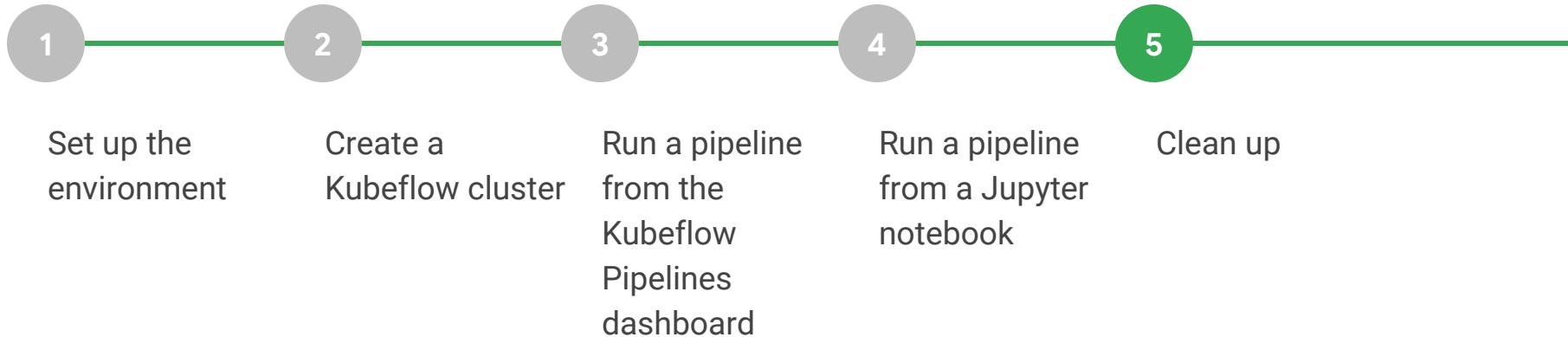
# Roadmap

- v0.4 RC out Dec. 17
- Better Jupyter Notebook integration
  - Spawner with PVC support
  - Fairing alpha
- TFJob & Pytorch operators in beta
- Application CRD
- Katib TFJob support
- Pipeline experiment comparison
- v1.0 Enterprise readiness
  - Model management
  - Hardened APIs
  - Clean deployments, upgrades
- **You tell us!** (Or better yet, help!)

Google Cloud

# Contributors

- Kubeflow is open
  - Open community
  - Open design
  - Open source
  - Open to ideas
- Get involved
  - github.com/kubeflow
  - kubeflow.slack.com
  - @kubeflow
  - kubeflow-discuss@googlegroups.com
  - Community call Tuesdays alternating 8:30am and 5:30pm Pacific
  - **Kubeflow Contributor Summit**
    - Q1 2019



`https://github.com/kubeflow/kubeflow`
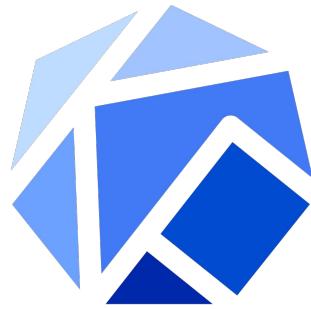
@amygdala @texasmichelle

# Questions?

Google Cloud

@amygdala @texasmichelle