E2E ML Platform on Kubernetes with just a few clicks

Mofi Rahman

http://tiny.cc/e2e-ml-k8s

ML and AI was a novelty

Data Scientists View of the World

Configuration

Data Collection

Data Verification

ML Code

Machine Resource Management

Analysis Tools

Feature Extraction

Process Management Tools

Serving Infrastructure

Monitoring

@moficodes
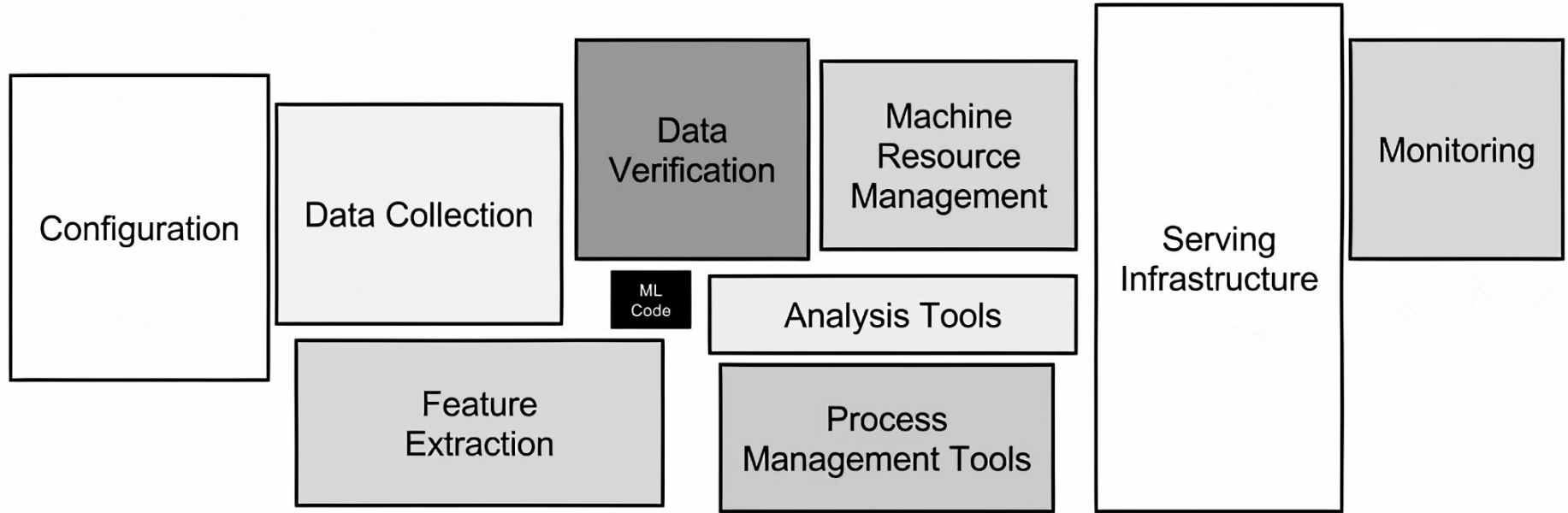
Now it is core part of every business

Data is everywhere

Understanding
this data is
more important
than ever

Rise of MLOps

What is MLOps

"the ability to apply DevOps principles to Machine Learning applications"
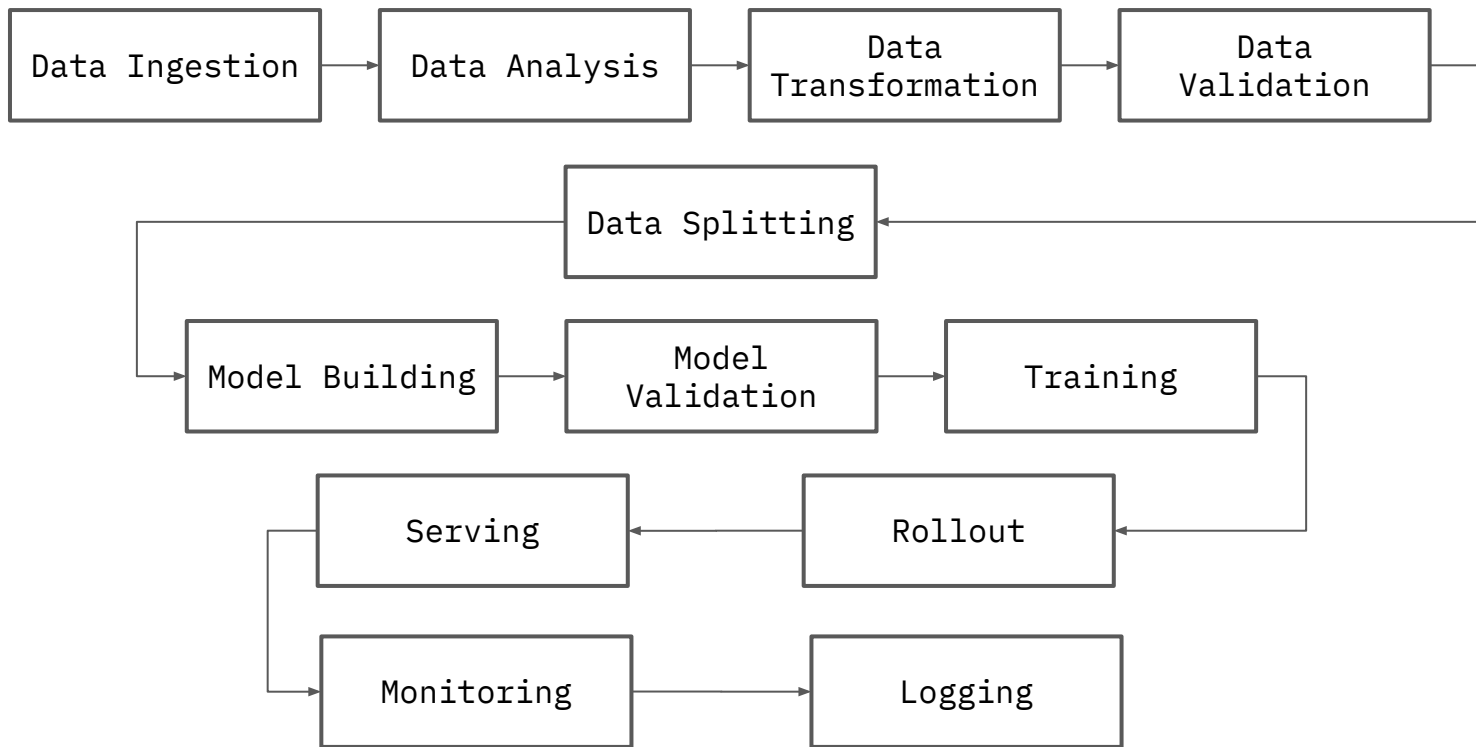-sig mlops cd foundation

Mofizur Rahman
Developer Advocate, IBM

Do Container Stuff,
Collect Stickers,
Write go code

@moficodes

E2E ML Platform

E2E ML Platform Features

```
┌──────────────┐     ┌──────────────┐     ┌──────────────┐     ┌──────────────┐
│Data Ingestion│ ──> │Data Analysis │ ──> │     Data     │ ──> │     Data     │
│              │     │              │     │Transformation│     │  Validation  │
└──────────────┘     └──────────────┘     └──────────────┘     └──────────────┘
```

Data Ingestion → Data Analysis → Data Transformation → Data Validation

Data Validation → Data Splitting

Data Splitting → Model Building → Model Validation → Training

Training → Rollout → Serving

Serving → Monitoring → Logging

@moficodes

Commercial E2E ML Platform Offerings

@moficodes

Most major cloud
provider has one

E2E ML Platform

- GCP AI Platform
- IBM Cloud Pak for Data & Watson Studio
- AWS Sage Maker
- Azure Machine Learning

…

And Many more

@moficodes

Pros

- Fully managed
- Works well with other cloud services
- Cloud scale
- Enterprise support

Cons

- Expensive
- Vendor Lock in
- Not easy to maintain dev env
- Not open source
- Code/Model are not portable

Why not DIY

DIY

- Uber
- Netflix
- Airbnb
- Lyft

    …

    And many more

Pros

- Full control over platform
- Owned by you, so no vendor lock in
- Customized to your needs

Cons

- Expensive
- You are on the hook
- Over time the platform becomes harder to maintain

E2E ML Platform wishlist

- Built on scalable infrastructure
- Uses existing tools data scientists already use
- Open source
- Supported by the industry
- Enterprise support options
- Portable

@moficodes

@moficodes

Kubeflow

Kubeflow

Open source project that contains a curated set of
tools/frameworks for ML workflows on Kubernetes

@moficodes

Kubeflow

Scalable
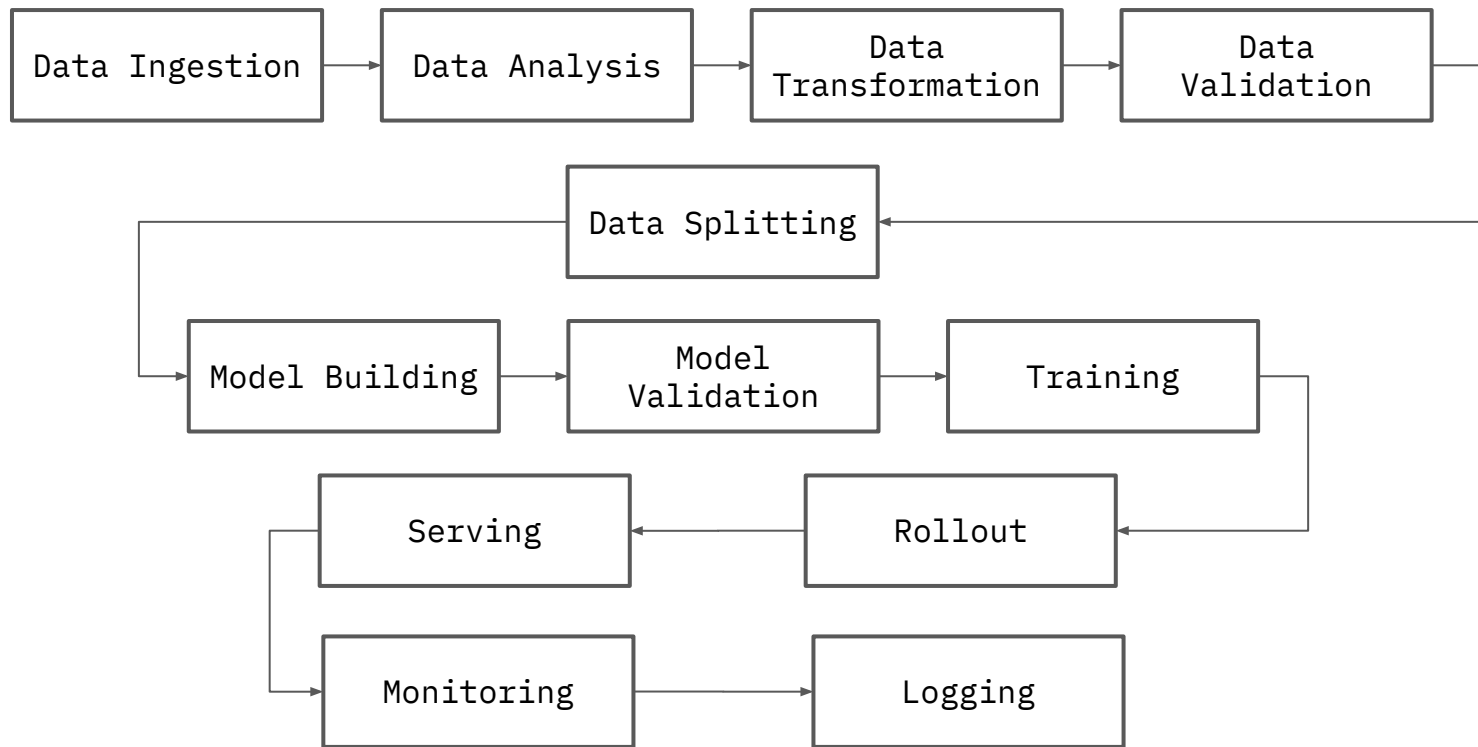
Composable

Portable

Open Source

Industry Supported

Multitenant

@moficodes

Scalable

- Built on top of Kubernetes
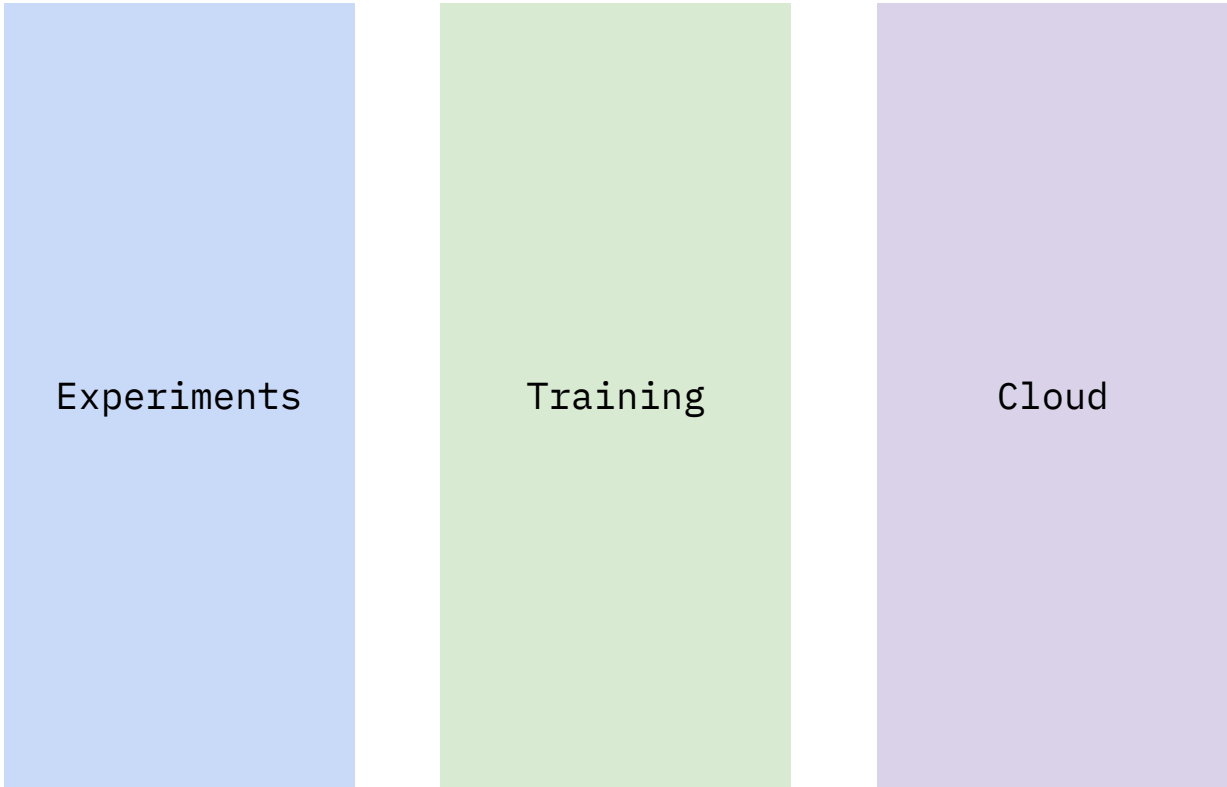- Scalability is built in
- Existing infrastructure skills

@moficodes

Composable

```
┌──────────────────┐     ┌──────────────────┐     ┌──────────────────┐     ┌──────────────────┐
│  Data Ingestion  │ ──▶ │  Data Analysis   │ ──▶ │      Data        │ ──▶ │      Data        │
│                  │     │                  │     │ Transformation   │     │   Validation     │
└──────────────────┘     └──────────────────┘     └──────────────────┘     └──────────────────┘
```

```
                        ┌──────────────────┐
                        │  Data Splitting  │ ◀────
                        └──────────────────┘
```

```
┌──────────────────┐     ┌──────────────────┐     ┌──────────────────┐
│  Model Building  │ ──▶ │      Model       │ ──▶ │    Training      │
│                  │     │   Validation     │     │                  │
└──────────────────┘     └──────────────────┘     └──────────────────┘
```

```
┌──────────────────┐     ┌──────────────────┐
│     Serving      │ ◀── │     Rollout      │ ◀────
└──────────────────┘     └──────────────────┘
```

```
┌──────────────────┐     ┌──────────────────┐
│    Monitoring    │ ──▶ │     Logging      │
└──────────────────┘     └──────────────────┘
```

@moficodes

Portable

- From local to training env
- From training env to cloud
- From cloud to cloud

Portability

Experiments

Training

Cloud

Portability (Reality)

Experiments

Training

Cloud

@moficodes

ML is no longer
a novelty

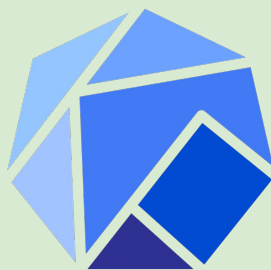# ML is no longer just research

ML is part of the core business

You quality control your software? Then you want to quality control your ML artifacts
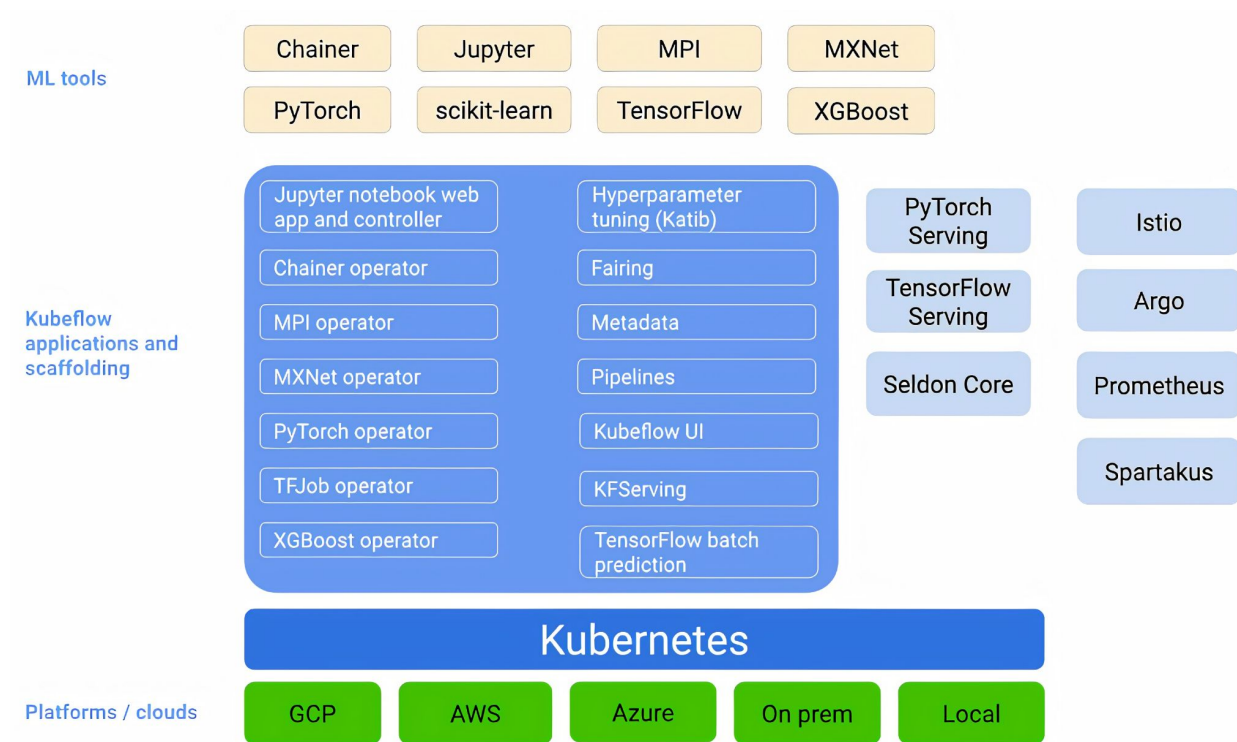
Portability

Experiments

Training

Cloud

@moficodes

Kubeflow Components

ML tools

Chainer	Jupyter	MPI	MXNet
PyTorch	scikit-learn	TensorFlow	XGBoost

Kubeflow applications and scaffolding

Jupyter notebook web app and controller	Hyperparameter tuning (Katib)
Chainer operator	Fairing
MPI operator	Metadata
MXNet operator	Pipelines
PyTorch operator	Kubeflow UI
TFJob operator	KFServing
XGBoost operator	TensorFlow batch prediction

PyTorch Serving

TensorFlow Serving

Seldon Core

Istio

Argo

Prometheus

Spartakus

Kubernetes

Platforms / clouds

GCP	AWS	Azure	On prem	Local

@moficodes

Kubeflow Components

- Dashboard
- Jupyter Notebook/Servers
- Machine Learning Framework
- Pipelines
- Serving
- Metadata and Base ML
- Feature Store
- Monitoring

Kubeflow Deployment

[Manifests repository](#) provides a catalog of Kubeflow application and common service manifests.

- kfdef + kfctl
- kustomize + kubectl
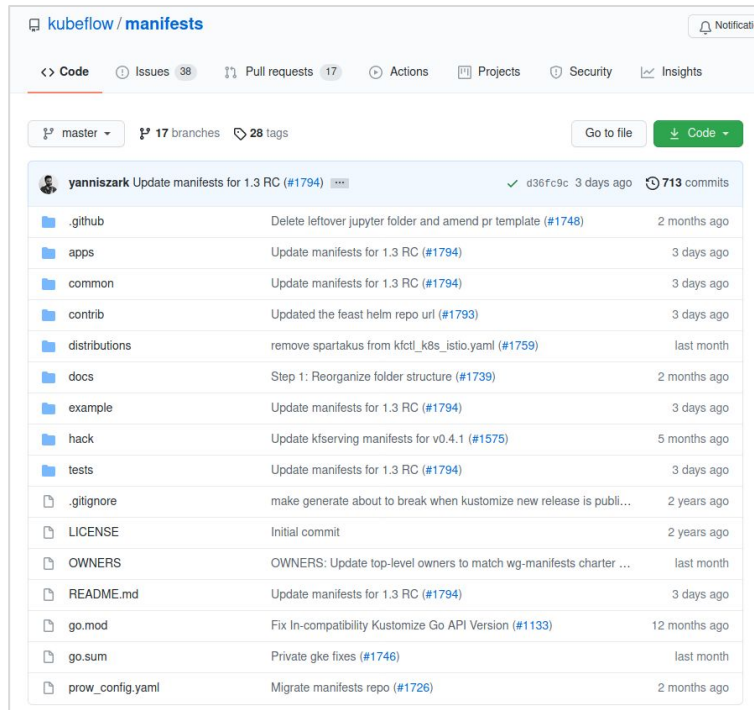
Manifests

- Manifests structure reorganized.
  - apps/, common/, contrib/
- Component manifests development moved to individual upstream app repos.
  - Previously, several versions of manifests existed.
  - Now only one source of truth.
  - Manifests repo copies component manifests at specific commit.
- Provides reference Kubeflow deployment.

@moficodes

Kubeflow Manifests Restructure

```
├── LICENSE
├── OWNERS
├── README.md
├── admission-webhook
├── application
├── argo
├── aws
├── cert-manager
├── common
├── default-install
├── dex-auth
├── docs
...                        1.2
├── experimental
├── gatekeeper
├── gcp
├── go.mod
├── go.sum
├── hack
├── istio
├── istio-1-3-1
├── jupyter
├── tektoncd
├── tests
├── tf-training
└── xgboost-job
```

```
├── LICENSE
├── OWNERS
├── README.md
├── apps
│   ├── admission-webhook
│       ...
│   └── xgboost-job
├── common
│   ├── cert-manager
│       ...
│   └── user-namespace
├── contrib
│   ├── application
│       ...                        1.3
│   └── tektoncd
├── distributions
│   ├── aws
│       ...
│   └── stacks
├── docs
├── example
│       └── kustomization.yaml
├── hack
└── tests
...
```

@moficodes

How This Can Help

Goals

- Improved accountability for maintaining component manifests.
- Increased modularity.
  - Pick and choose individual components easier.
- Smoother deployment experience.

Another Deployment Option

Operators

Operators

- Methods of packaging, deploying, and managing a Kubernetes application.
- Represent human operational knowledge.
- Custom control loop using CRDs.
- Implement and automate:
  - Day-1 (installation, configuration, etc.)
  - Day-2 (re-configuration, failover, update, etc.)

@moficodes
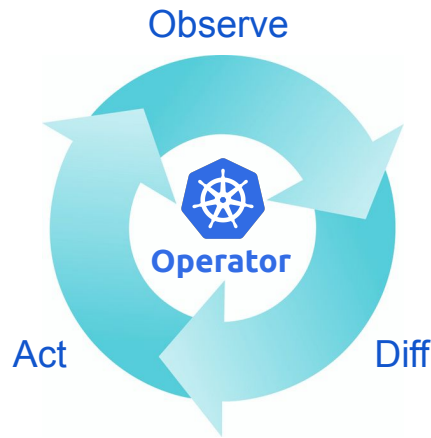
Kubeflow Operator

- Deploy, monitor, and manage the lifecycle of Kubeflow.
- Built using Operator Framework.
- Available on OperatorHub.io.
- More info: kubeflow.org/docs/operator

@moficodes
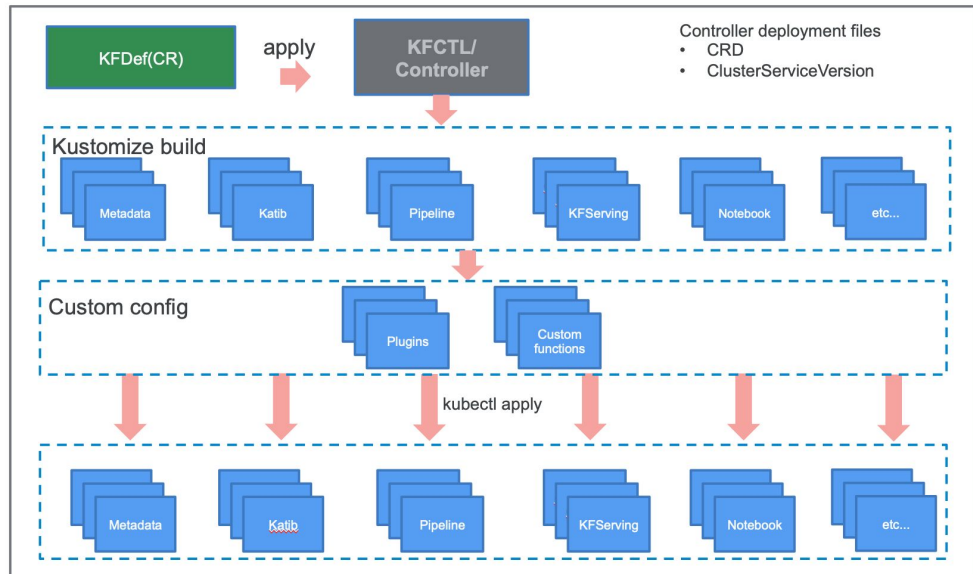
It's not all
sunshine and
rainbows

Kubeflow is growing rapidly

So is the use
cases and the
underlying
components

Challenge

- Kubeflow has many moving parts
- Each component has their own release lifecycle
- Each component has their own update/upgrade path
- Interoperability is not always guaranteed
- It is all built on top of open-source technology which is also changing rapidly
- Each platform has minor differences that add up

@moficodes

The Future

- Kubeflow 1.3 will bring many needed updates that should help resolve many of these challenges.
- The official installation will rely on kubectl and kustomize.
- Manifest repo is being restructured so it is easier to navigate.

@moficodes

References

- Manifests Repository: github.com/kubeflow/manifests
- Community Links: kubeflow.org/docs/about/community
- Operator Framework: operatorframework.io