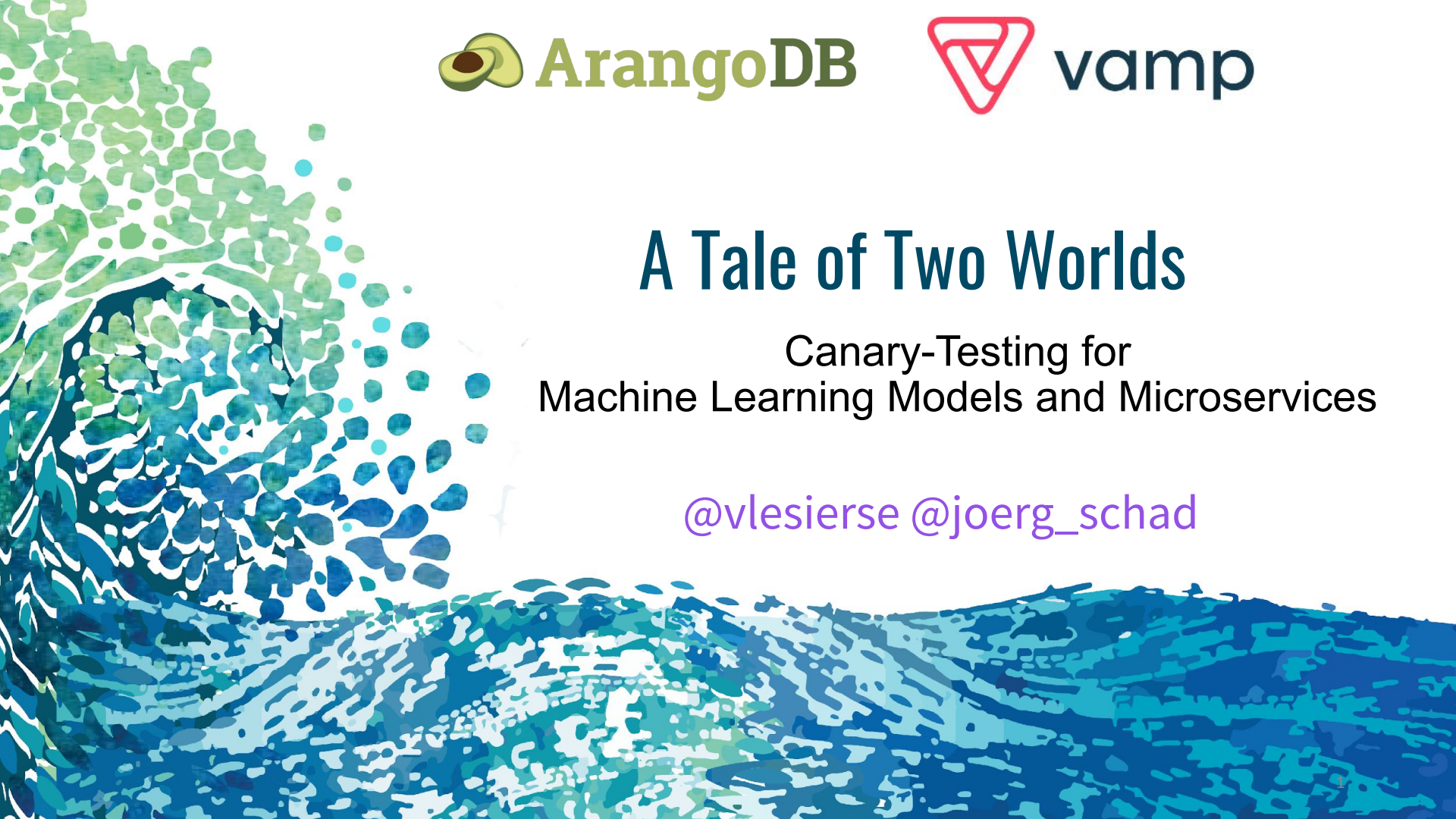

ArangoDB
vamp
A Tale of Two Worlds
Canary-Testing for
Machine Learning Models and Microservices
@vlesierse @joerg_schad

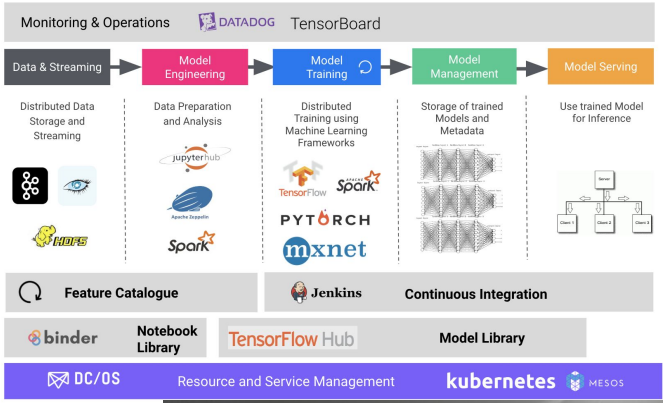
Vincent Lesierse

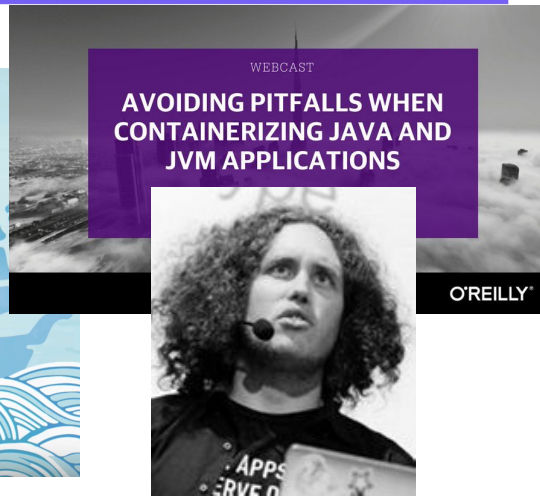
Technical Product Manager - Vamp.io

- **Previous**
  - **Software Engineer**
  - **Software Architect**
  - **Cloud Architect**
- **@vlesierse**

Jörg Schad

Head of Engineering & Machine Learning @ ArangoDB





- Previous
  - Suki.ai
  - Mesosphere
  - PhD Distributed DB Systems

- @joerg_schad

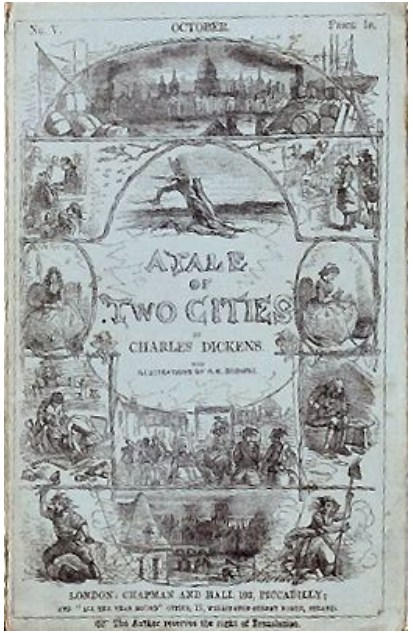
A tale of two worlds

It was the best of times,
where Microservices allowed to break code in manageable pieces,

It was the worst of times
where we actually had to deploy all those pieces,

It was the age of wisdom,
where Machine Learning allowed for great discoveries,

it was the age of foolishness,
where we suffered to productionize these models...



A tale of two worlds

Deployment of Microservices

Deployment of Machine Learning Models

Microservices...

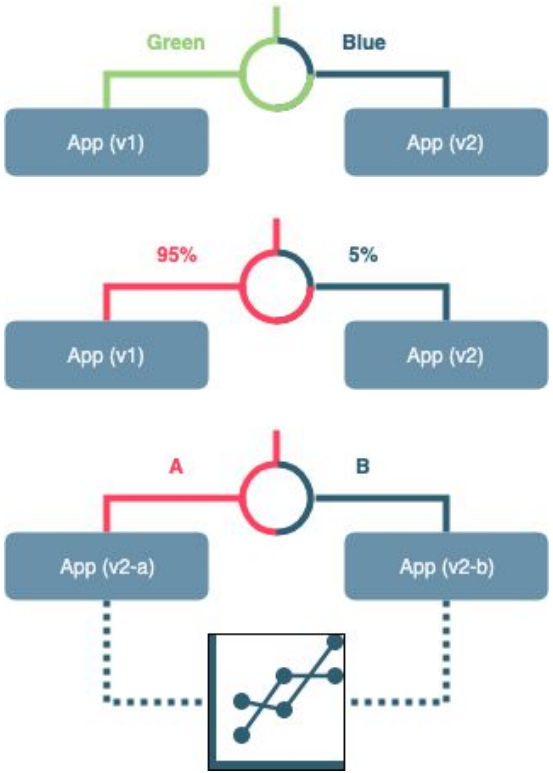
- Independent teams owning their services
- More frequent deployments
- Not more frequent incidents
- Enabling experimentation

Microservice Deployments

Blue/Green

Canary

A/B Testing
Shadow Deployments



Our mission: Provide an easy-to-use and powerful solution to transform software from ideas into value as efficiently as possible, using cloud-native technologies.

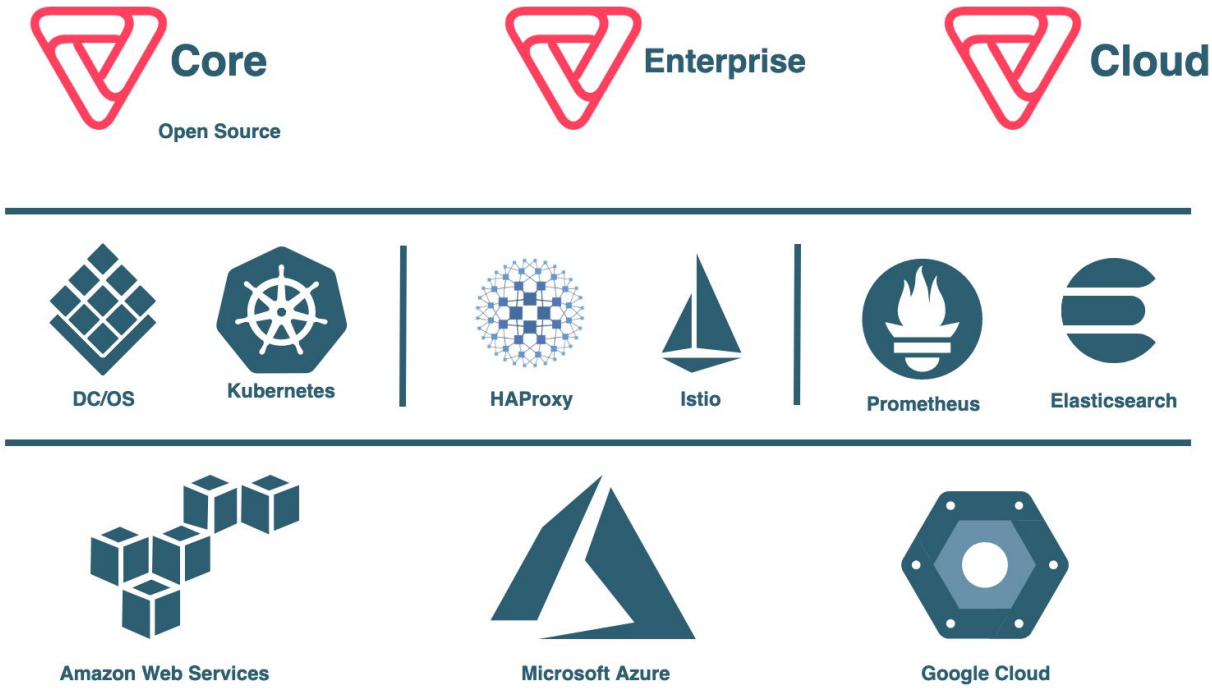
Important: Releasing is (much) more than technical deployment. Continuous validation requires observability and “golden metrics”. Processes are crucial. Tools and technologies are means to an end.

2014: canary testing & releasing and SLA-based auto-scaling on top of Mesos+Marathon, using HAProxy for networking (service-discovery/load-balancing/mesh/ingress)

2016: added Kubernetes support

Kubecon 2019: added Istio support

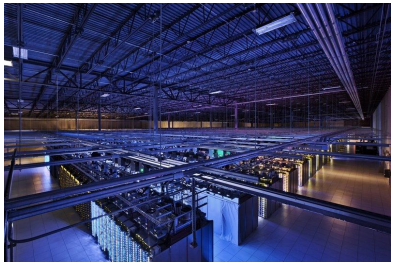
Vamp Stack

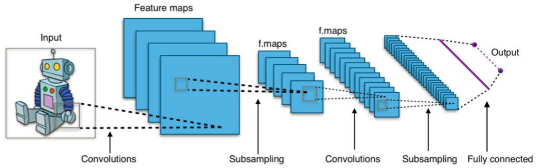

Core
Open Source
Enterprise
Cloud
DC/OS
Kubernetes
HAProxy
Istio
Prometheus
Elasticsearch
Amazon Web Services
Microsoft Azure
Google Cloud

Machine Learning Model Serving….

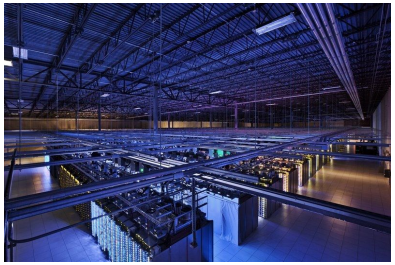
- Productionize trained models
- Retraining and redeployment
- Large number models
- Metadata based Policies
- Testing with live data

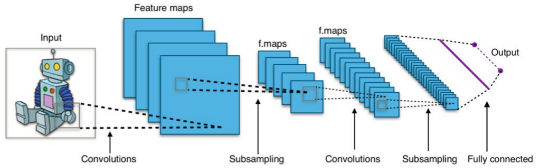

Science
Projects
Data
Free
Python
Sas
Big
Top
Lists
Use
Datascience
Dataset
Best
Excel
Part
Here
Gather
Websites
Big
Here



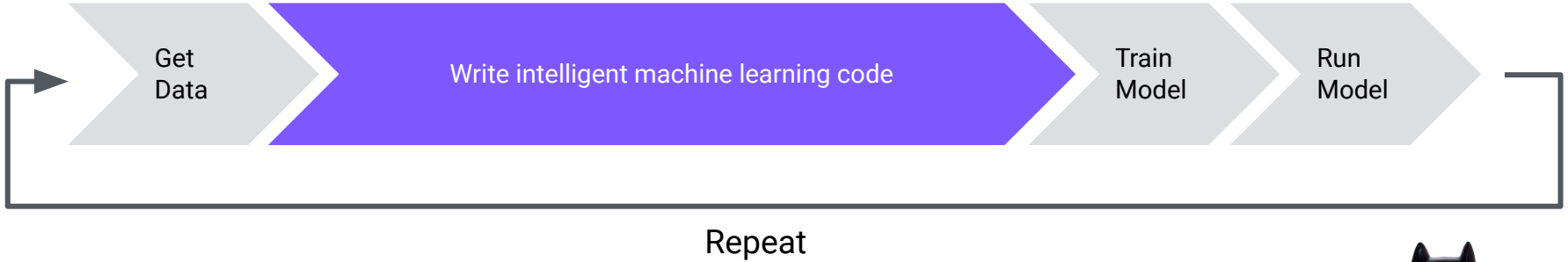

Feature maps
Input
f.maps
f.maps
Output
Convolutions
Subsampling
Convolutions
Subsampling
Fully connected


Science
Projects
Data
Free
Python
Sas
Big
Top
Lists
Use
Datascience
Best
Excel
Part
Dataset
Gather
Websites
Here

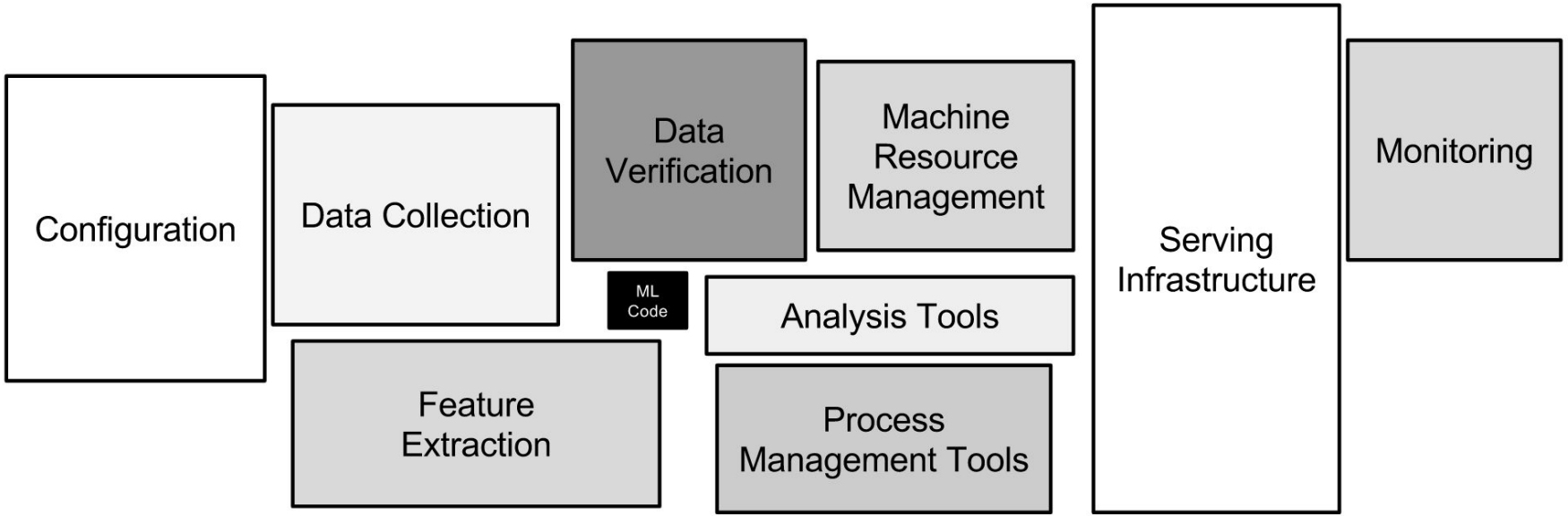



Input
Feature maps
f.maps
f.maps
Output
Convolutions
Subsampling
Convolutions
Subsampling
Fully connected

What you want to be doing



What you're actually doing



Sculley, D., Holt, G., Golovin, D. et al. Hidden Technical Debt in Machine Learning Systems

Monitoring & Operations

DATADOG

TensorBoard

Data & Streaming → Model Engineering → Model Training → Model Management → Model Serving

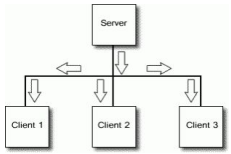
Distributed Data Storage and Streaming

Data Preparation and Analysis

Distributed Training using Machine Learning Frameworks

Storage of trained Models and Metadata

Use trained Model for Inference



Feature Catalogue

Jenkins

Continuous Integration

binder

Notebook Library

TensorFlow Hub

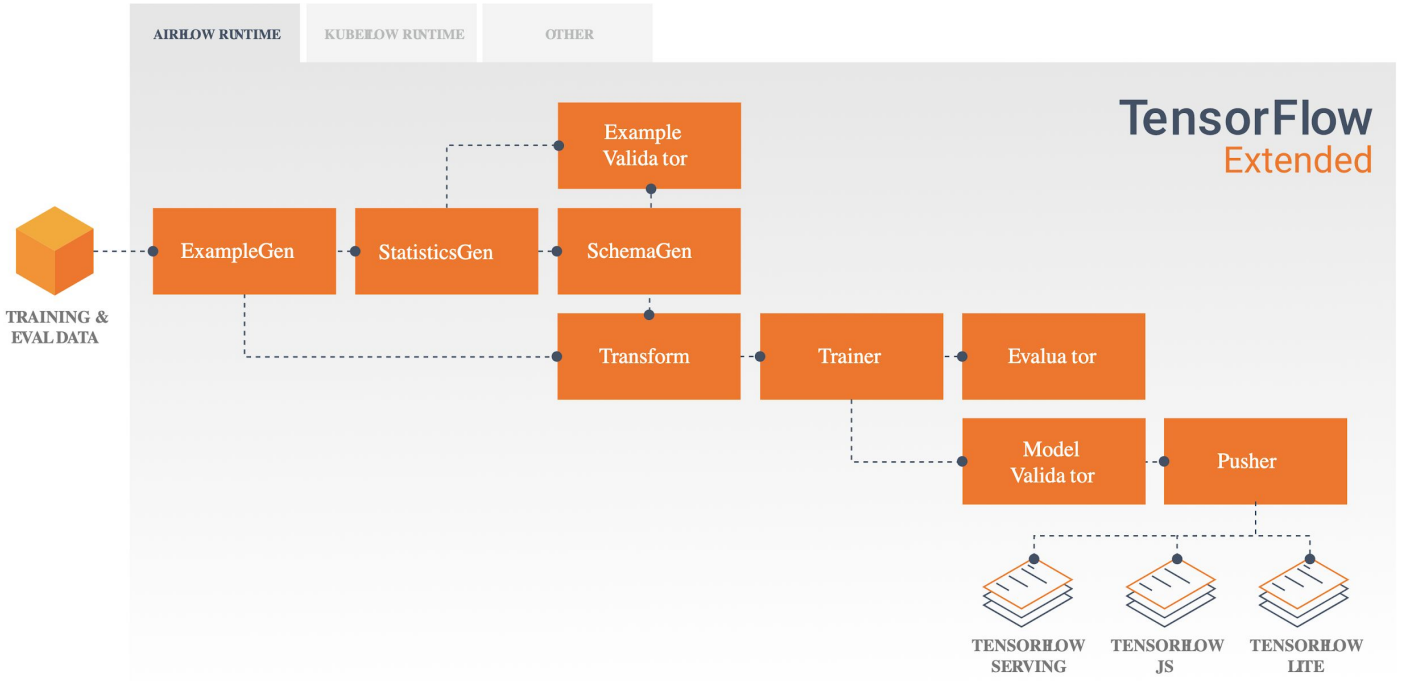
Model Library

DC/OS

Resource and Service Management

kubernetes

TensorFlow Extended



Challenge: Persona(s)

The Rise of the *DataOps Engineer*

Combines two key skills:

- Data science
- Distributed systems engineering

The equivalent of *DevOps* for *Data Science*

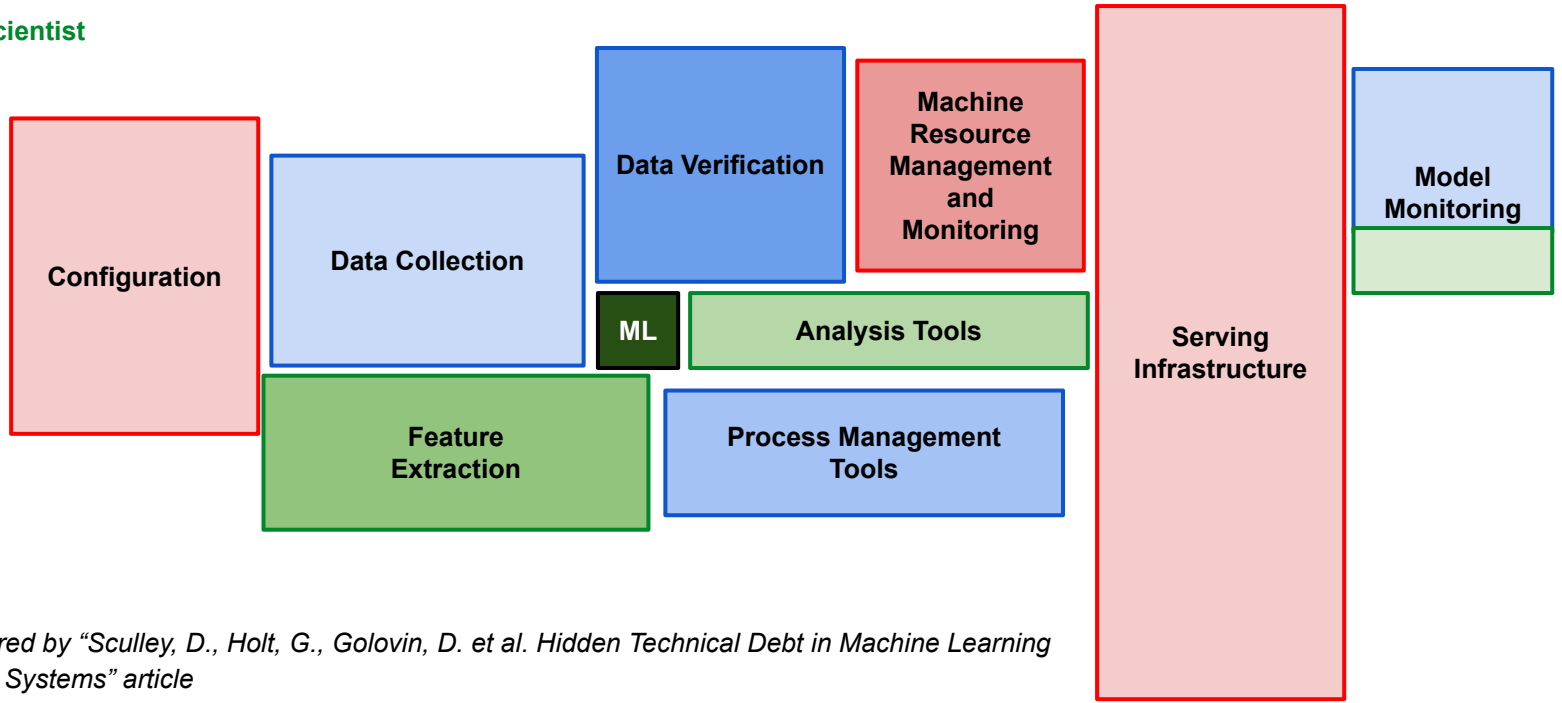
- **Build** automation software to run machine learning systems
- **Operate** systems so they're available, scalable, and performant
- **Evangelize** tools and best practices among data scientists

Division of Labor

System Admin/ DevOps

Data Engineer/DataOps

Data Scientist



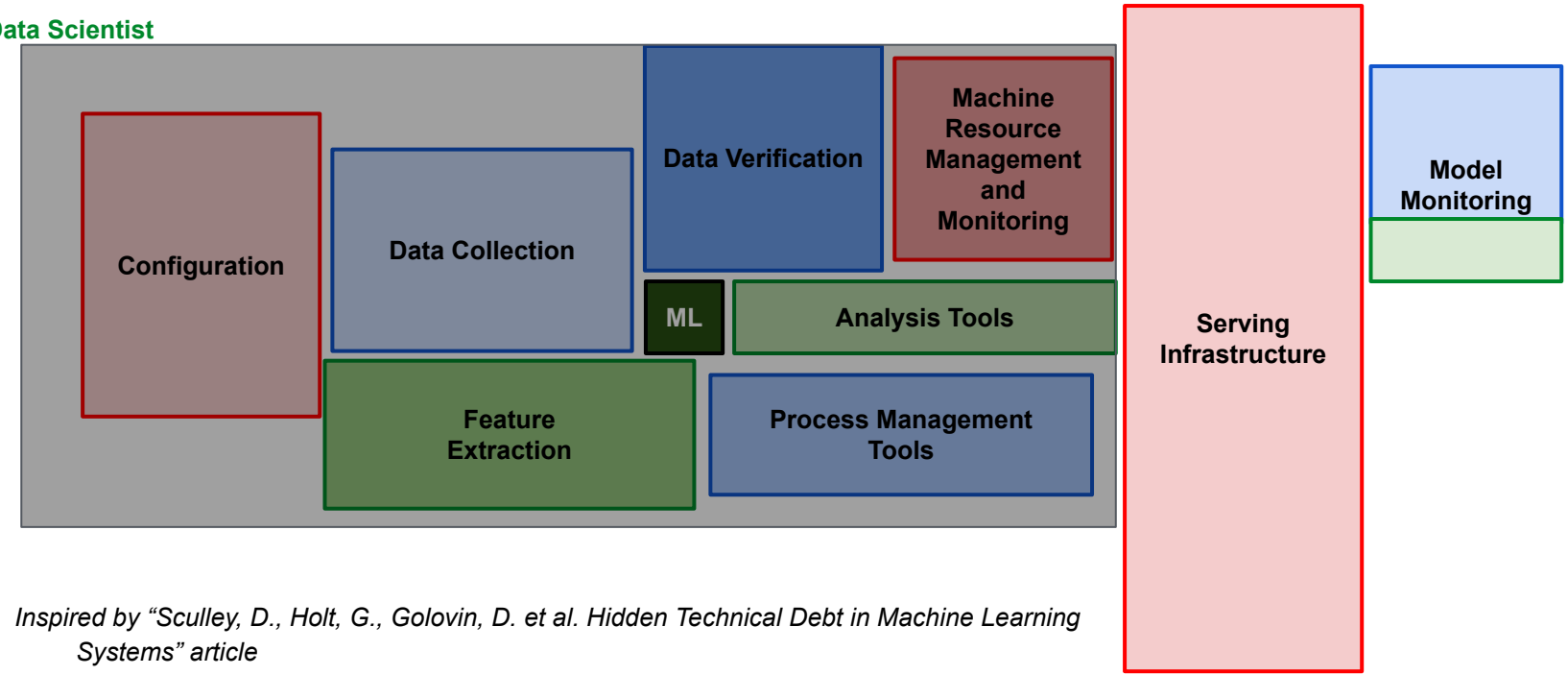
Inspired by "Sculley, D., Holt, G., Golovin, D. et al. Hidden Technical Debt in Machine Learning Systems" article

Division of Labor

System Admin/ DevOps

Data Engineer/DataOps

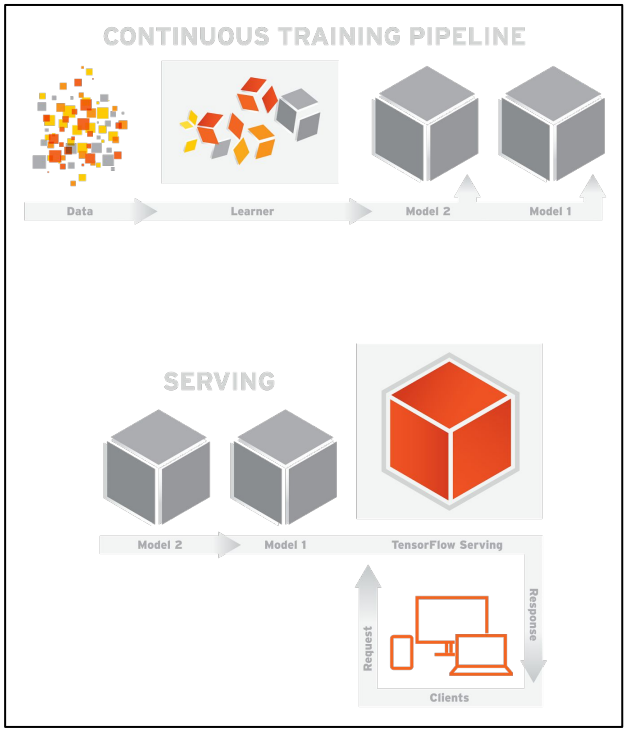
Data Scientist



Inspired by "Sculley, D., Holt, G., Golovin, D. et al. Hidden Technical Debt in Machine Learning Systems" article

Machine Learning Model Serving

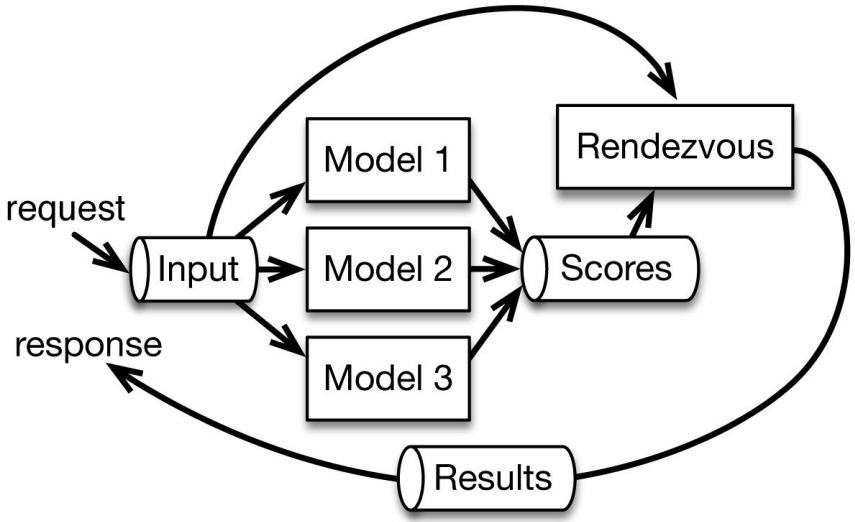
- Deploying models
  - Choice…
  - Metrics
- Updating models
  - Zero downtime
  - Target environment
- Testing models
  - Test model with live data
- Ensemble Decision
  - Multiple models working together



https://www.tensorflow.org/tfx/guide/serving

Machine Learning Model Serving

- Deploying models
  - Choice…
  - Metrics
- Updating models
  - Zero downtime
  - Target environment
- Testing models
  - Test model with live data
- Ensemble Decision
  - Multiple models working together



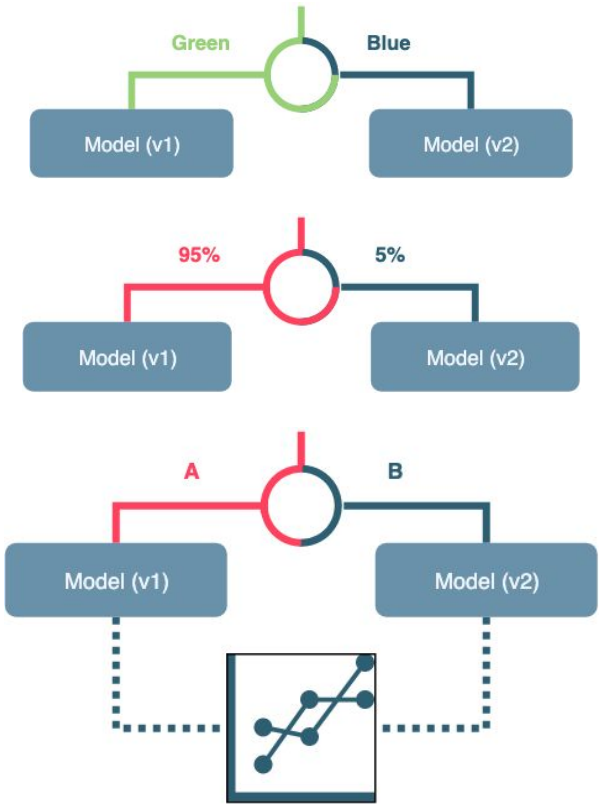
https://mapr.com/ebooks/machine-learning-logistics

Machine Learning Model Deployments

Updating model

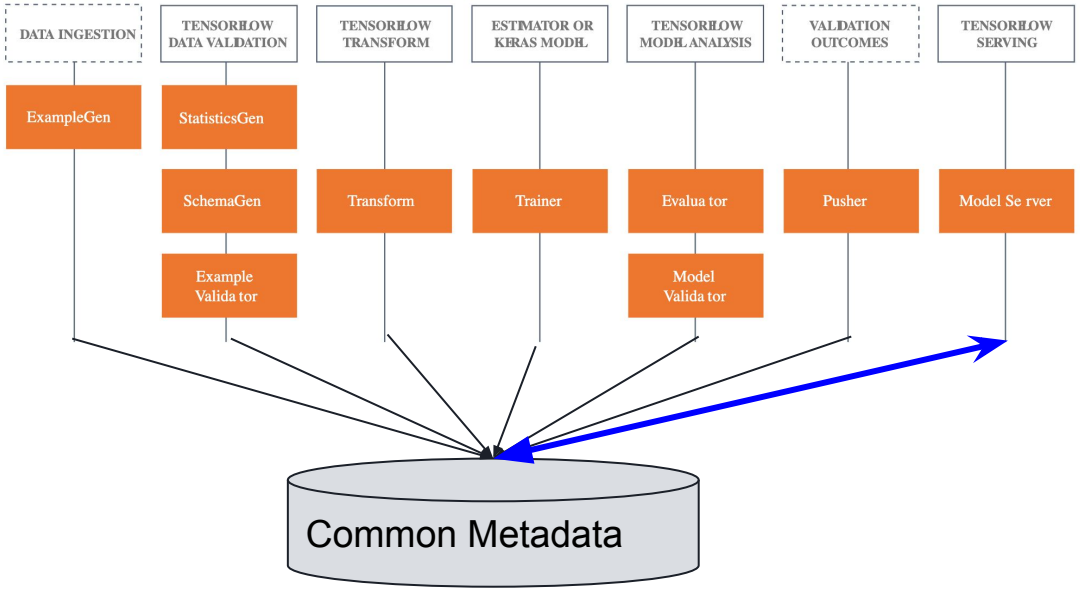
Testing model

Ensemble/
Testing a model

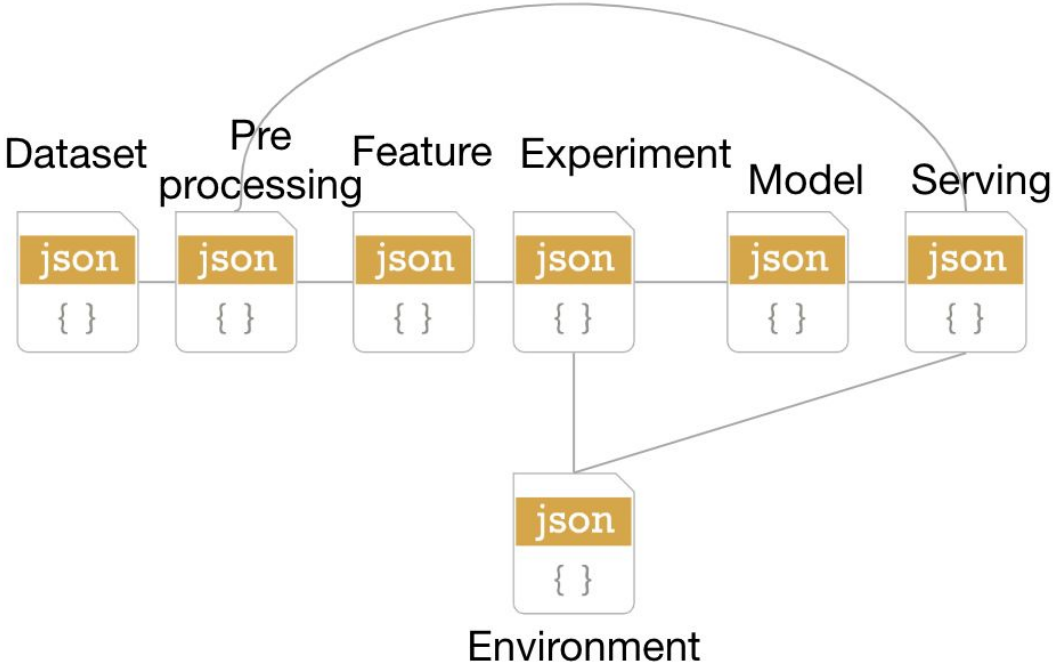


Metadata… Which model to pick?

- Accuracy
  - Which…
- Latency
- Environments
- Data Privacy
- ….

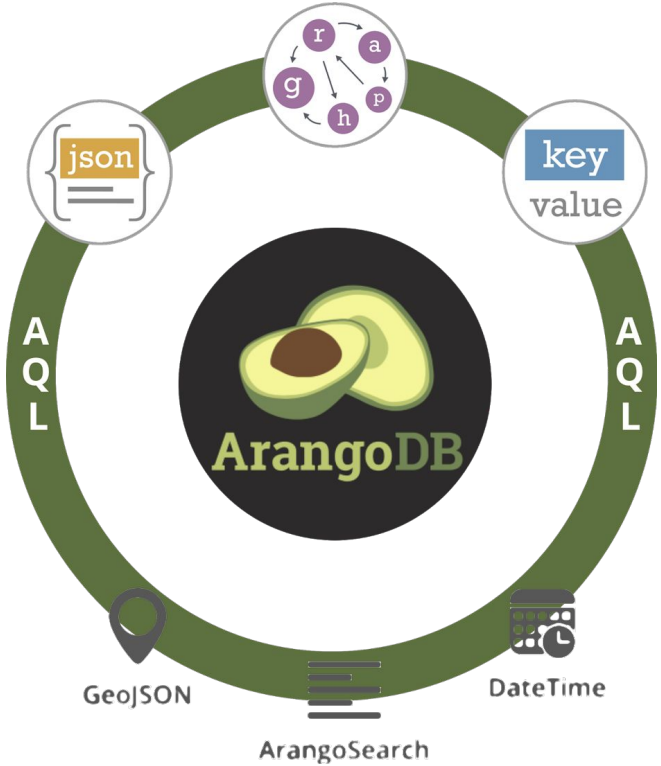


Metadata… How to store...



ArangoDB

- Native Multi Model Database
  - Stores, K/V, Documents & Graphs
- Distributed
  - Graphs can span multiple nodes
- AQL - SQL-like multi-model query language
- ACID Transactions including Multi Collection Transactions

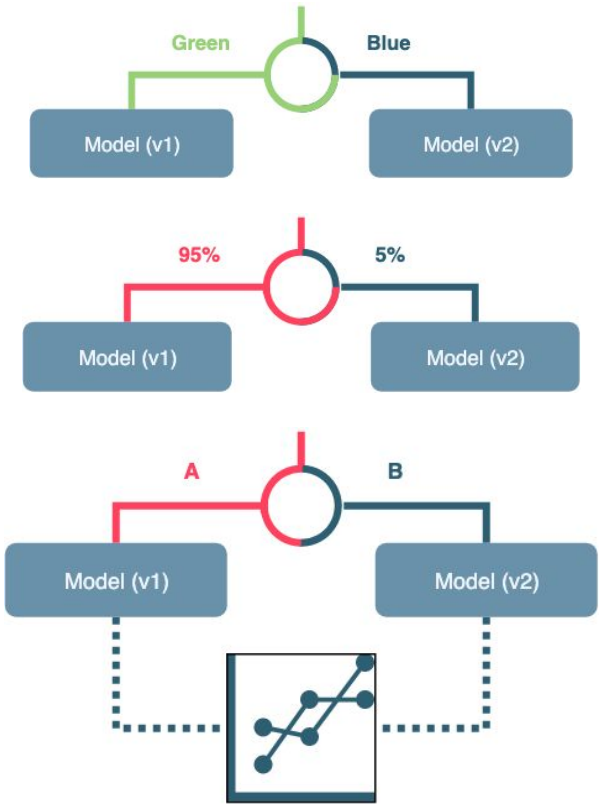


Machine Learning Model Deployments

Updating model

Testing model

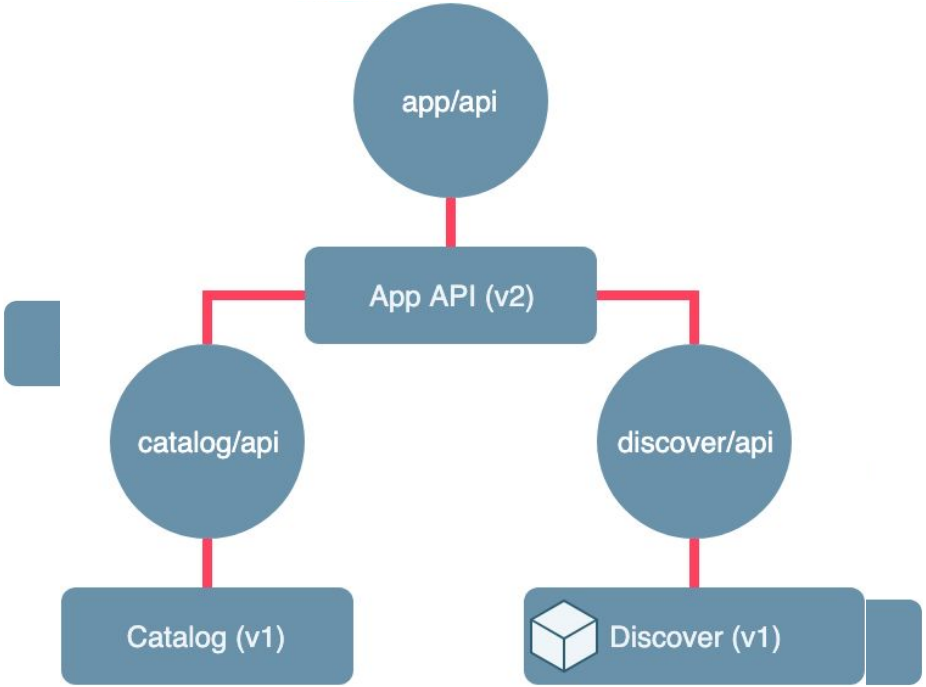
Ensemble/
Testing a model



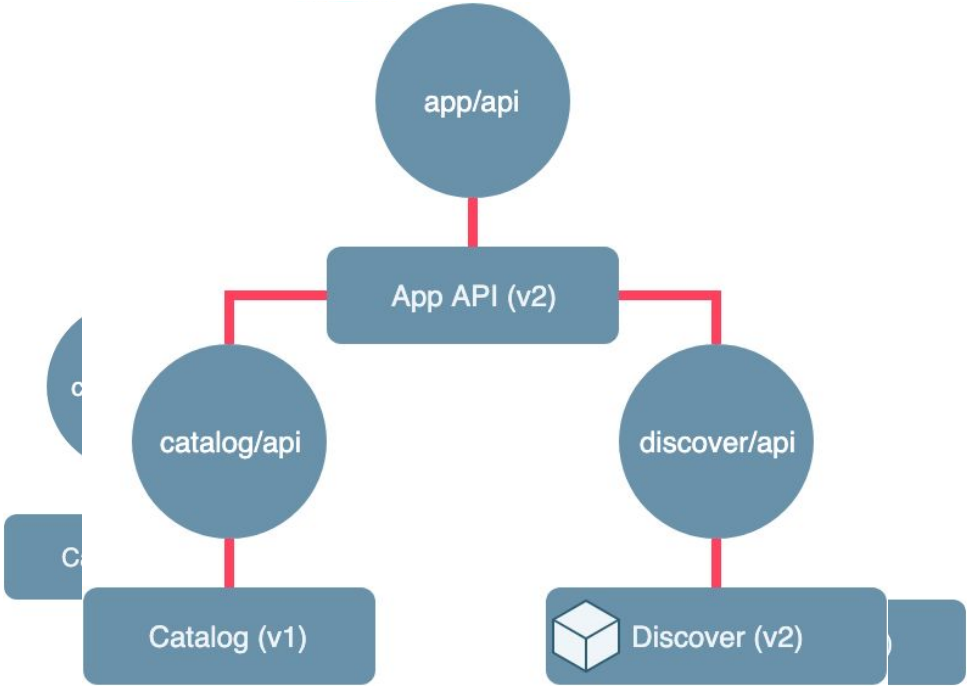
A tale of ~~two~~ one world

Deployment of Services

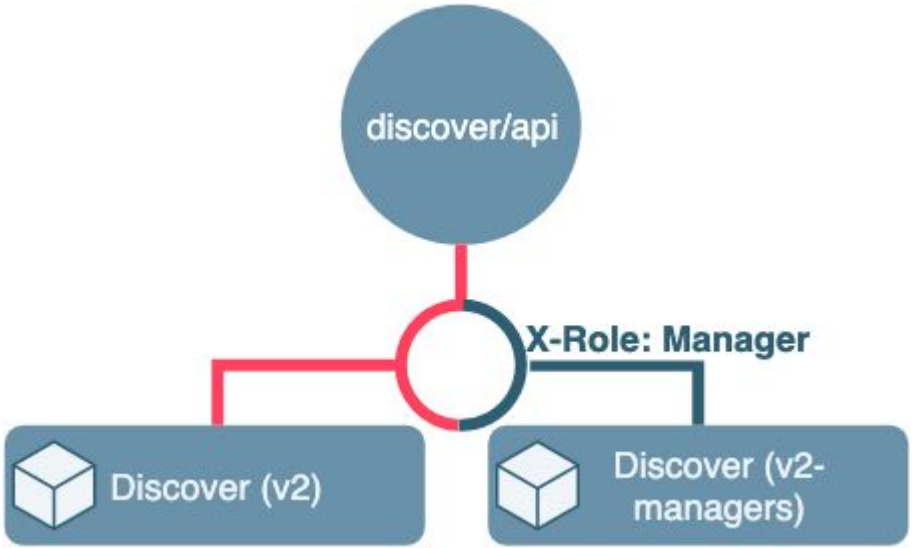
Demo: Canary Release a Service



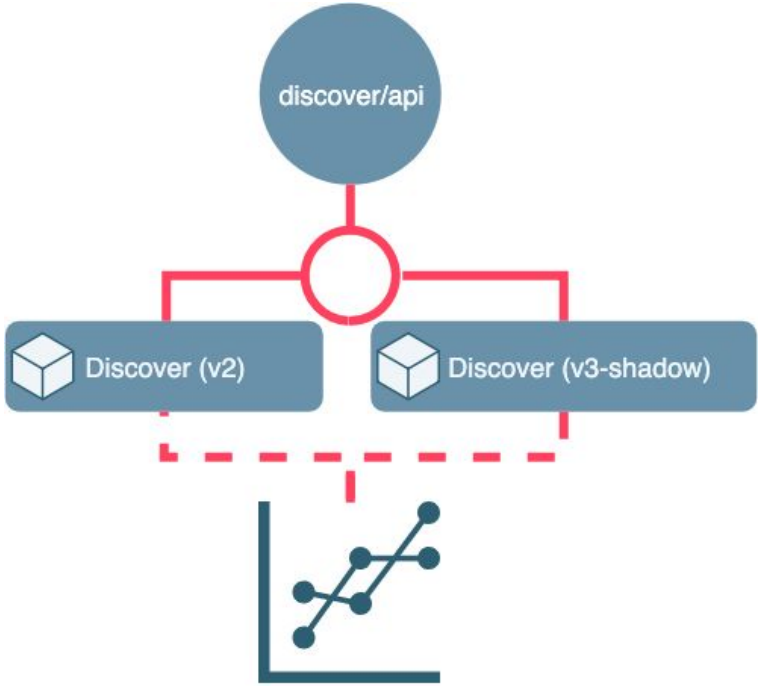
Demo: Canary Release a Model



Scenario: Model Segmentation



Scenario: Shadow Traffic



What’s next for Vamp?

- Integrating Istio for smart networking
- AI/ML based workflows & policies for self-learning and optimisation
- More CI and APM integrations
- Additional release, validation, experimentation and optimisation workflows
- Hybrid containers & serverless support
- Cloud-based version with new & improved core “engine”

Engage with us

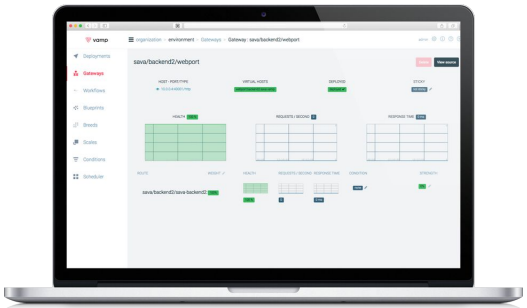
- Talk to us at Mesosphere stand P13
- Take a look at Vamp: https://github.com/magneticio/vamp
- Istio preview: https://vamp.io/vamp-istio-preview

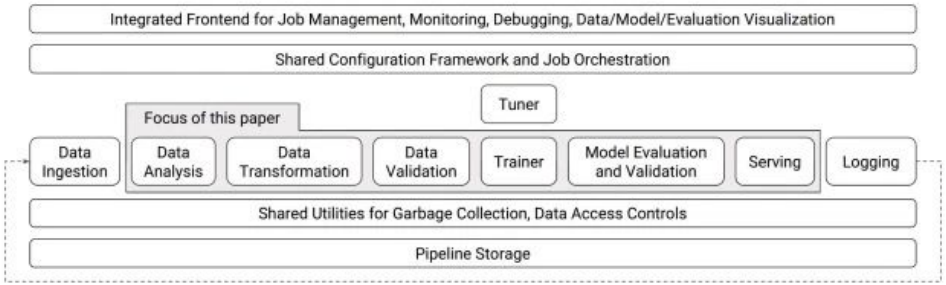
- @arangodb
- Take a look at ArangoDB:
- https://www.arangodb.com/
- ArangoML Preview: https://github.com/arangoml/arangopipe

Download the demo: https://github.com/magneticio/smartcon

@vlesierse @joerg_schad

A Tale of Two Worlds
Canary-Testing for Both ML Models and Microservices





vamp

ArangoDB