BIZREACH

ビズリーチの機械学習基盤

株式会社ビズリーチ システム本部 AI室
機械学習エンジニア
右田 尚人

2019年2月23日
第33回 Machine Learning 15minutes!

自己紹介

BIZREACH

- 名前: 右田 尚人（ミギタ ナオト）
- 2015/04〜2018/5: 富士通
  - ニューラルネットワーク系の自然言語処理の研究開発
  - AIサービスの導入支援・PoC（FAQ検索など）
- 2018/7〜: ビズリーチ
  - レコメンドエンジン開発
  - 機械学習基盤開発

ビズリーチでの機械学習活用

企業と求職者の**マッチング**、求人の職種・業種・年収**推定**
など様々な箇所で機械学習を活用

例: <u>求人内容から職種・業種を推定</u>

課題

- 様々な求人サイトから収集しているため横断的な職種・業種がない

内容

- 求人情報から自然言語処理で求人の特徴を抽出
- その特徴を学習して職種業種を推定
- 検索条件やデータ分析などで利用

参考: BizReach Tech Blog
https://tech.bizreach.co.jp/posts/276/gtc2018-japan-report/

なぜ私たちは機械学習基盤を必要としたのか？

BIZREACH

①データサイエンティスト・機械学習エンジニアの作業負荷低減

例: 新たなレコメンドエンジンを実装する時

前に作ったバッチ処理と同じ形で作っておいてね
deployの手順書はこれを参照してね
既存の動作についてちょっとだけ説明するね。

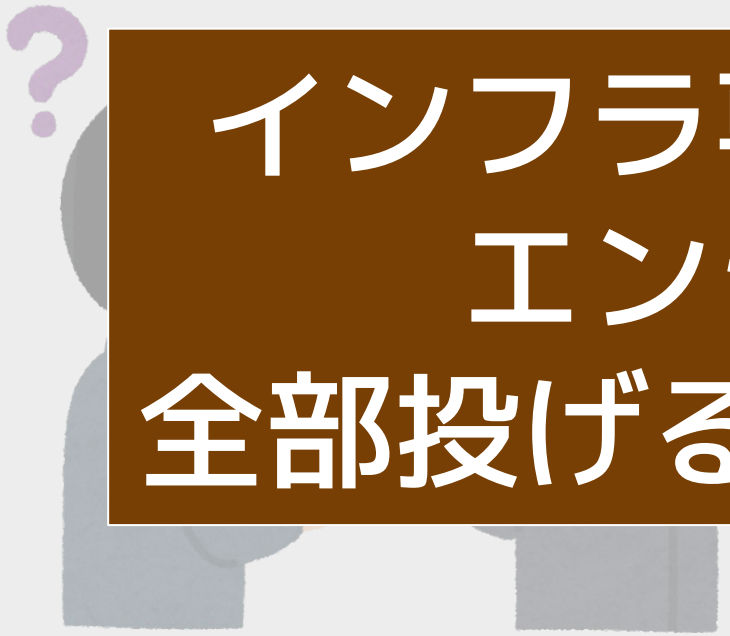
バッチ処理の動作は**Docker**のImageとしてBuildされてます。処理を動かす時は、AWSの**Lambda**をKickしてね。処理が始まると、**AutoScalingGroup**の設定を変更してスポットインスタンスを払い出して、そのインスタンス上で処理が実施されるよ。実施中のログは**CloudWatch**に吐かれていて、もしErrorが発生したら、Slackに通知が飛ぶようになってる。ここまでの設定は、**CloudFormation**を使うことで適宜定義しているよ。
あ、そうそう、データについては事業部側の**S3**から取ってくるので専用の**Role**を用意して事業部側に権限を付与してもらう必要もあるね。
まあ、これ読めばわかるから、あとよろしく！

なぜ私たちは機械学習基盤を必要としたのか？

BIZREAC月

①データサイエンティスト・機械学習エンジニアの作業負荷低減

例: 新たなレコメンドエンジンを実装する時



前に作ったバッチ処理と同じ形で作っておいてね
deployの手順書はこれを参照してね
ね。

インフラ専門じゃない
エンジニアに
全部投げるの無理だから

てBuildされ
odaをKickし
oupの設定を
して、その
実施中のロ
rrorが発生
る。ここま
での設定は、**CloudFormation**を使うことで適宜定義しているよ。
あ、そうそう、データについては事業部側の**S3**から取ってくるので専用の**Role**を用意して事業部側に権限を付与してもらう必要もあるね。
まあ、これ読めばわかるから、あとよろしく！

なぜ私たちは機械学習基盤を必要としたのか？

②インフラ側の特殊な条件

- 様々な環境に対して基盤を展開する必要アリ
- 事業部側がサービスを管理しているため サービス x (本番|開発)環境ごとにAWSアカウントを発行
- 今後は他クラウドサービス（GCP、Azure）対応の可能性も

まとめ

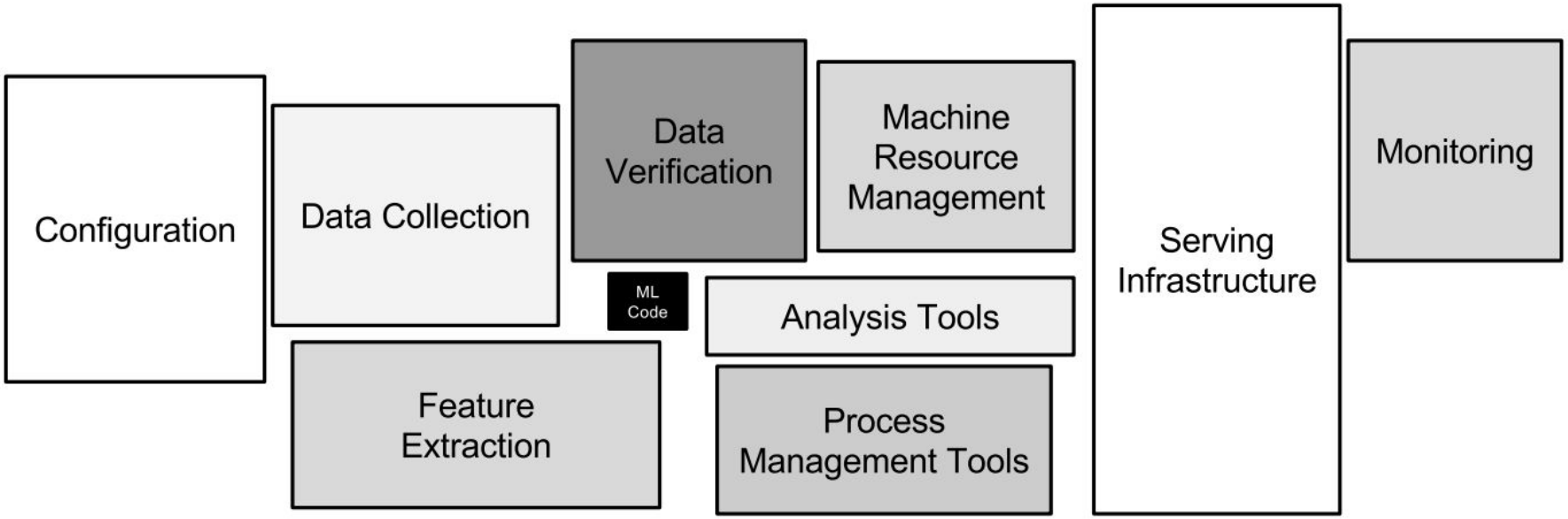
複数の環境に**お手軽に展開**できて、
インフラを**あまり意識しない**で使える基盤が欲しい

から作りました

BIZREACH

機械学習サービスの課題

Hidden Technical Debt in Machine Learning Systems NIPS2015

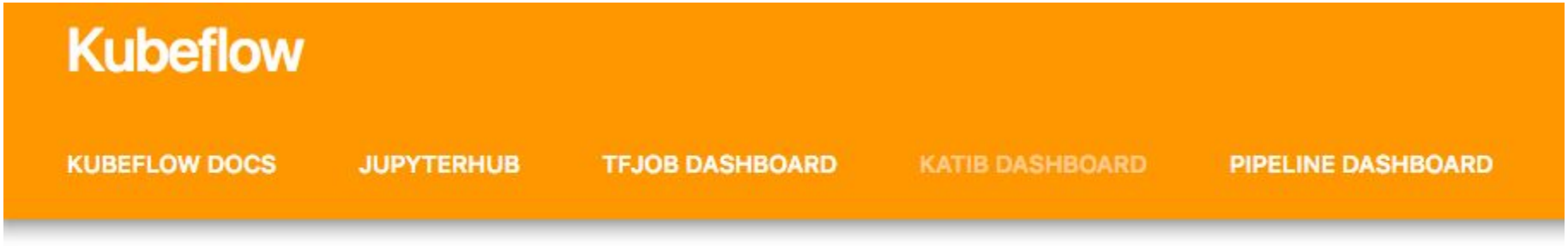


https://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf

世の中の機械学習基盤

BIZREACH

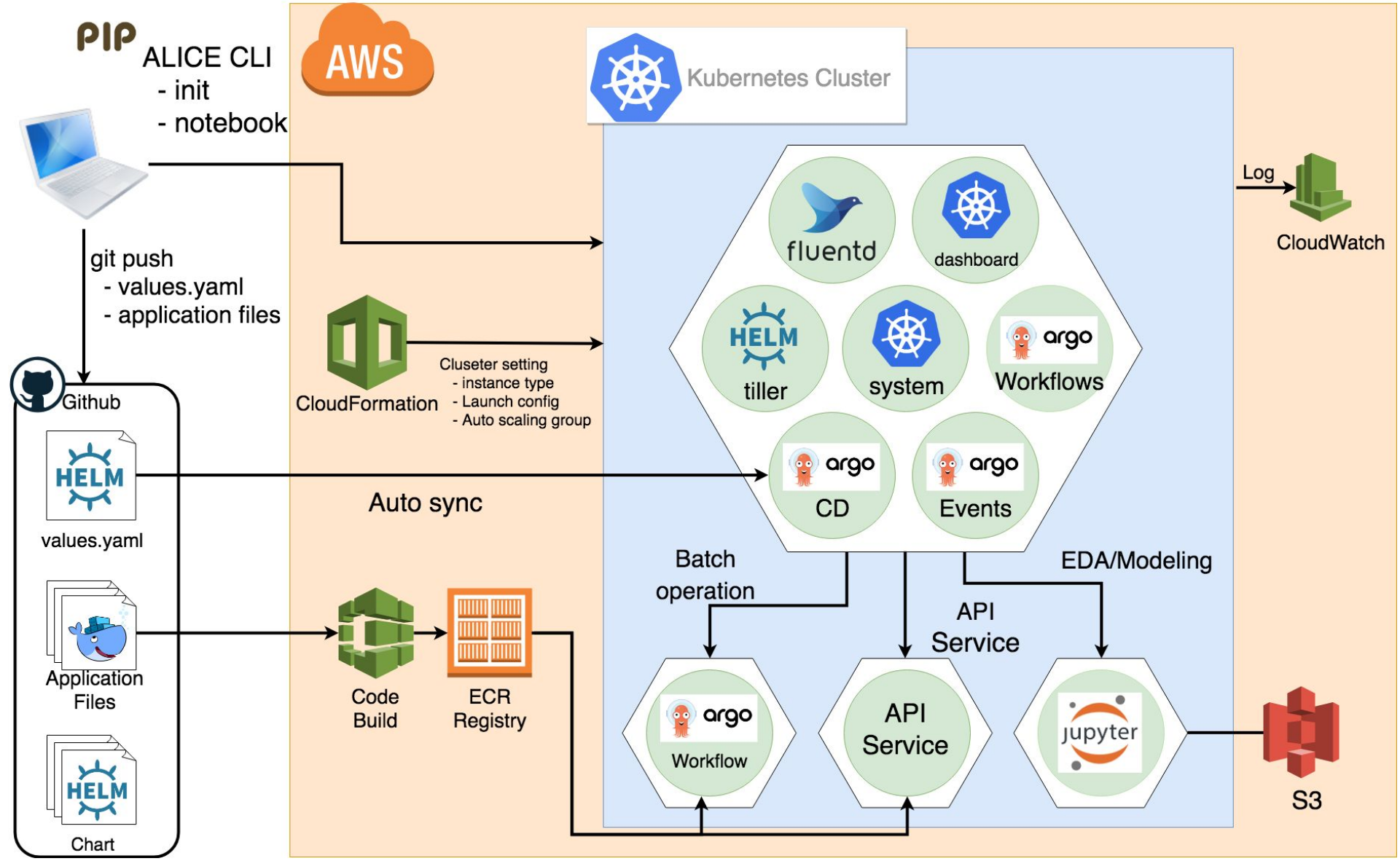
- KubeFlow
  - KubeCon2017年でgoogleが発表したOSSプロジェクト
  - kubernetes上に機械学習基盤を構築
  - 現状だと発展途上かつGCP依存が多そう
    （PipelinesはGKE依存）



https://www.kubeflow.org/docs/other-guides/accessing-uis/

ビズリーチの機械学習基盤 ALICE

BIZREACH



values.yaml 概要

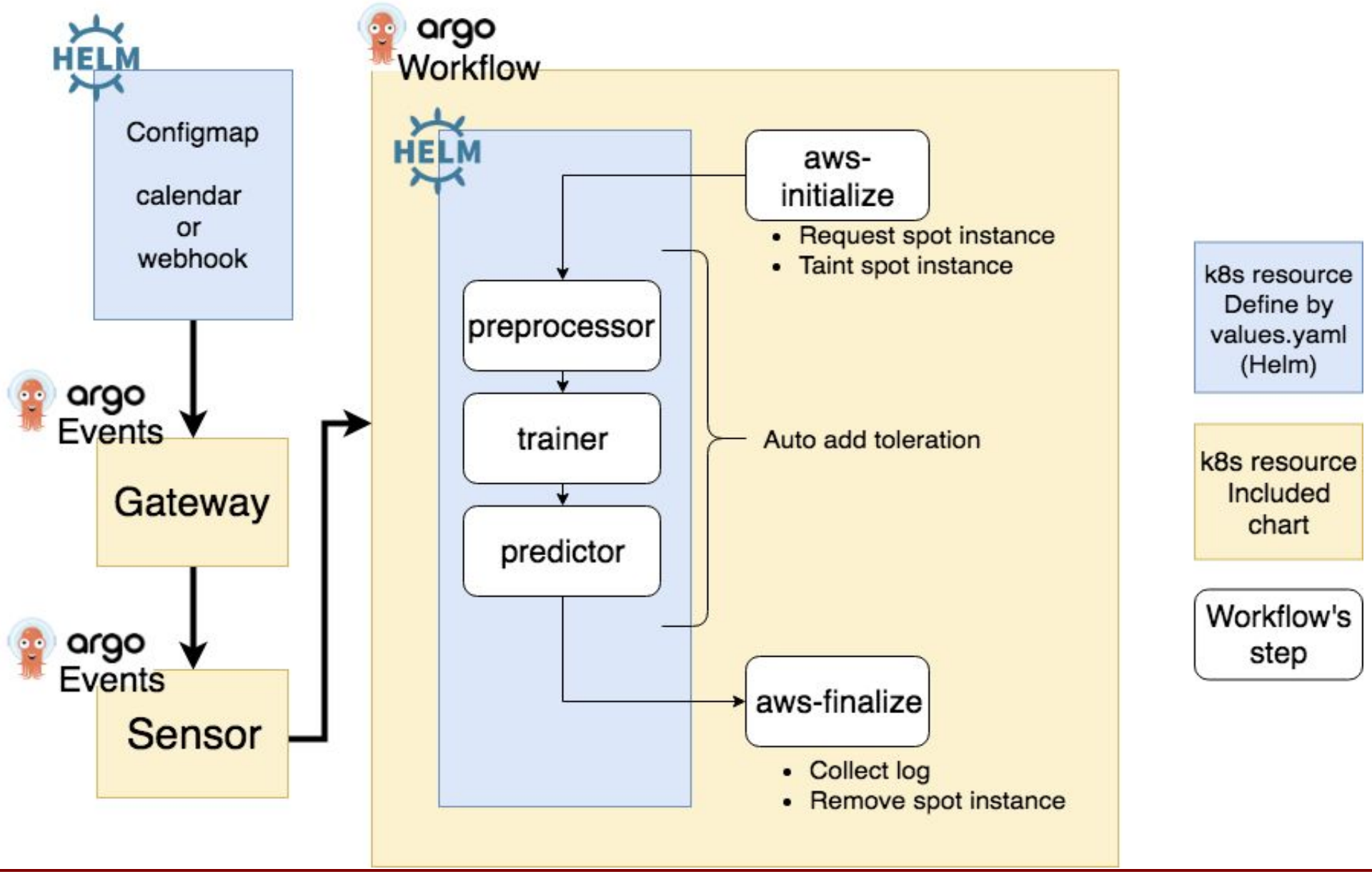
```yaml
##############################################
# Environment settings
##############################################

# environment. local or aws are available.
env: aws

# aws settings
aws:
  instanceType: t3.small
  spotPrice: "0.008"
  volumeSize: "8"

argocd: false

##############################################
# Deployment settings
##############################################

# deployment environment
# dev, stg, prd
depEnv: dev

# workflow or webhook or calendar
mode: workflow

# only valid for webhook mode
webhook:
  port: "12000"

# only valid for calendar mode
# you must specify "schedule" or "interval" when you use calendar mode.
# "schedule" takes precedence over "interval".
calendar:
  schedule: "*/30 * * * *"
  interval: 30m

##############################################
# Application settings
##############################################

# argo workflow
containerSteps:
- - name: preprocess
    template: preprocessor
- - name: training
    template: trainer
- - name: prediction
    template: predictor

# container definitions.
containerTemplates:
- name: preprocessor
  container:
    image: "XXX.ecr.amazonaws.com/sample/preprocessor:latest"
- name: trainer
  container:
    image: "XXX.ecr.amazonaws.com/sample/sample/trainer:latest"
- name: predictor
  container:
    image: "XXX.ecr.amazonaws.com/sample/sample/predictor:latest"
```

worker Nodeのスペック
- ボリューム
- インスタンスタイプ
- スポットインスタンスの入札価　　など

インフラに依らないデプロイ仕様の設定
(アプリを**どのように動かしたい**か)
- 動作モード: calendar/webhook
- モード固有設定

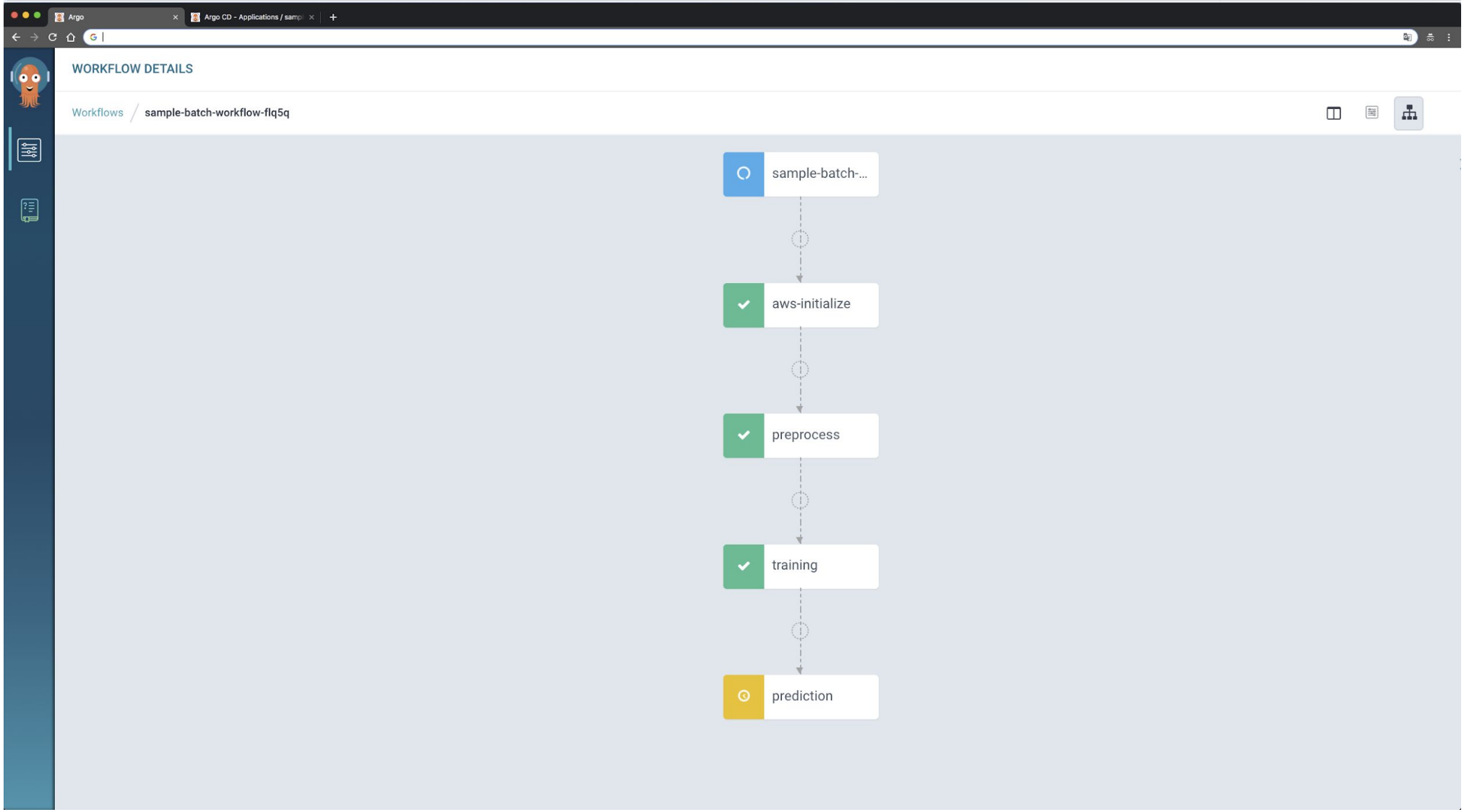
アプリケーション動作設定
- Workflow実行順
- 起動コンテナ定義

Chart概要



Workflowの見え方

BIZREAC月



ビズリーチの機械学習基盤 ALICE まとめ

BIZREACH

複数の環境にお手軽に展開できてインフラをあまり意識しないで使える基盤をkubernetesを使って開発

- ALICE CLIの提供
  - init: kubernetesクラスタをコマンド1つで展開
  - notebook: jupyter notebookをクラスタ上に展開
- アプリDeploy時の設定は、1枚のファイルを書くだけ
  - 最低限のAWS, Argoの知識があればDeploy可能になった
  - 設定ファイルの変更は、Argo CDでAuto sync。GitOpsを実現
- Argoによって柔軟なworkflowを定義可能
  - preprocess、training、predictionなどのジョブごとにコンテナやマシンスペックを選べる

今後

- 脱AWS依存
  - CodeBuild -> skaffold?
  - fluentd + CloudWatch Logs -> Prometheus?
- インフラエンジニアとの連携
  - Kubernetes周りは登場人物も多くてキャッチアップが大変
  - 専門の人に頼りたい
- ALICEを外部公開
  - まずはhelm chartから公開予定


BIZREACH