

Building Scalable Data Pipelines With Argo Workflows



spectrum labs

David Joyce

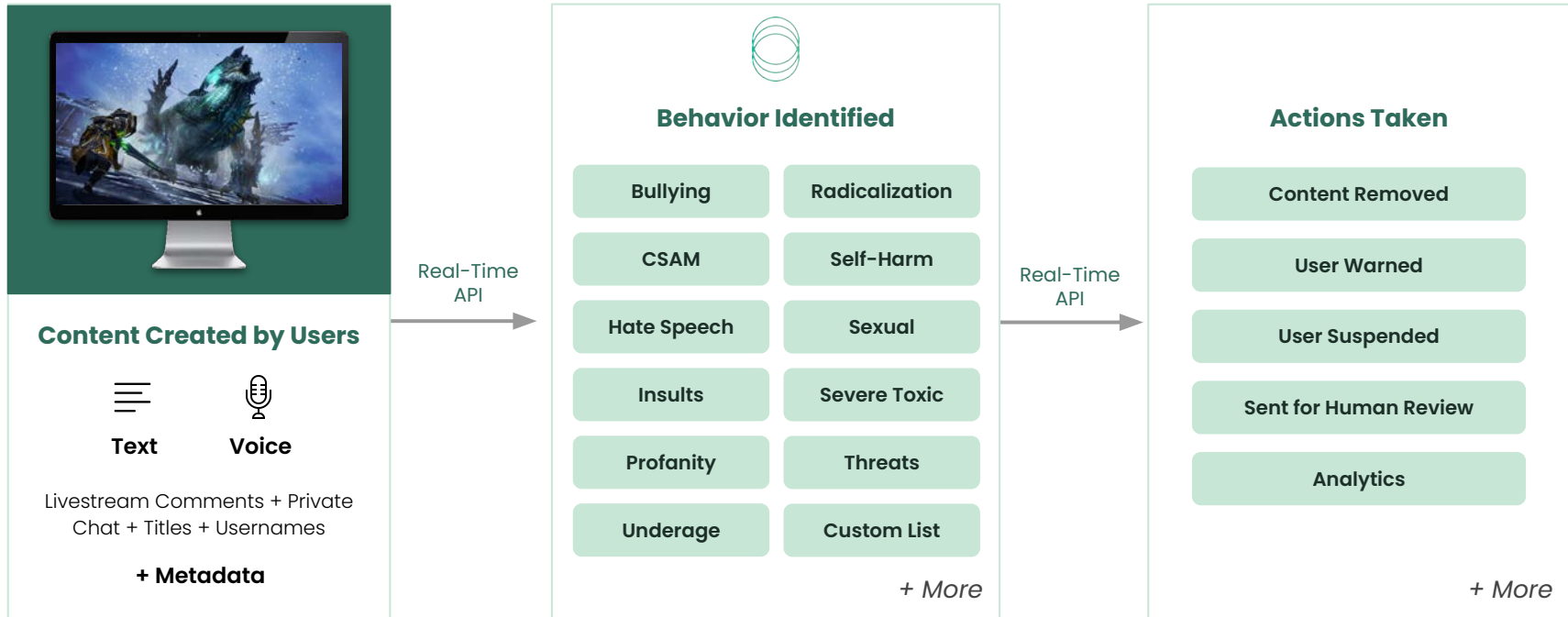
Principal Data Engineer

Contents

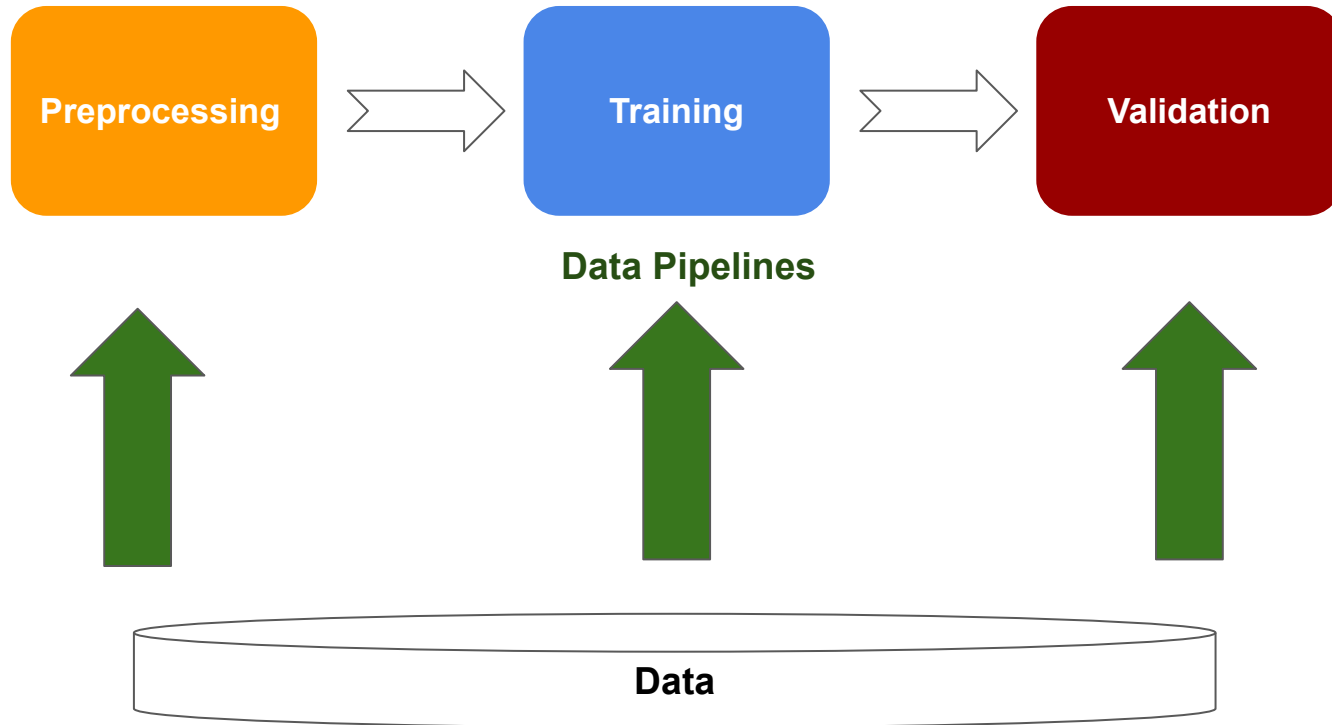
1. **Argo introduction + demo**
2. **Our pipelines + demo**
3. **Our Argo SDK**
4. **Argo events + demo**
5. **Argo challenges**
6. **Q&A / Discussion**



Trust & Safety Technology



Supporting Data Science Process



Argo Workflows

Workflows

Events

Continuous Deployments

Rollouts

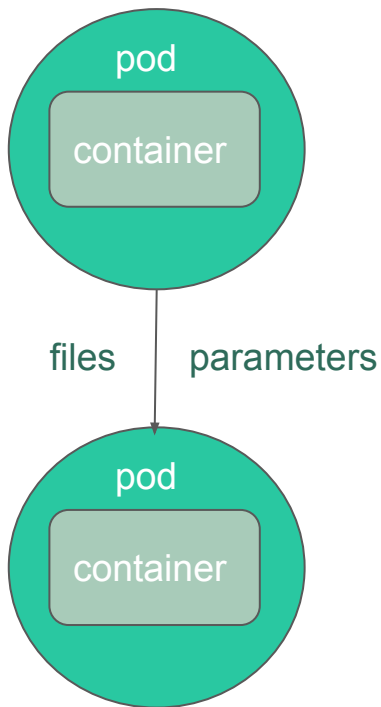


“Argo Workflows is an open source container-native workflow engine for orchestrating parallel jobs on Kubernetes”

Argo on



kubernetes



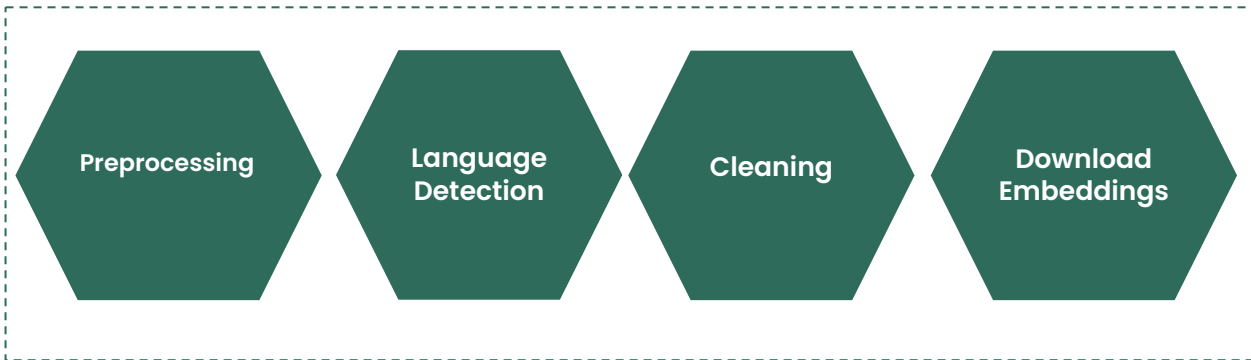
```
dag:
  tasks:
    - name: file-to-artifact-step
      template: file-to-artifact
      arguments:
        parameters:
          - name: path
            value: "{{inputs.parameters.path}}"
    - name: file-output-step
      dependencies: [file-to-artifact-step]
      template: file-output
      arguments:
        artifacts:
          - name: message
            from: "{{tasks.file-to-artifact-step.outputs.artifacts.message}}"
```

Argo Basics Demo



Our Data Pipelines

Iroh





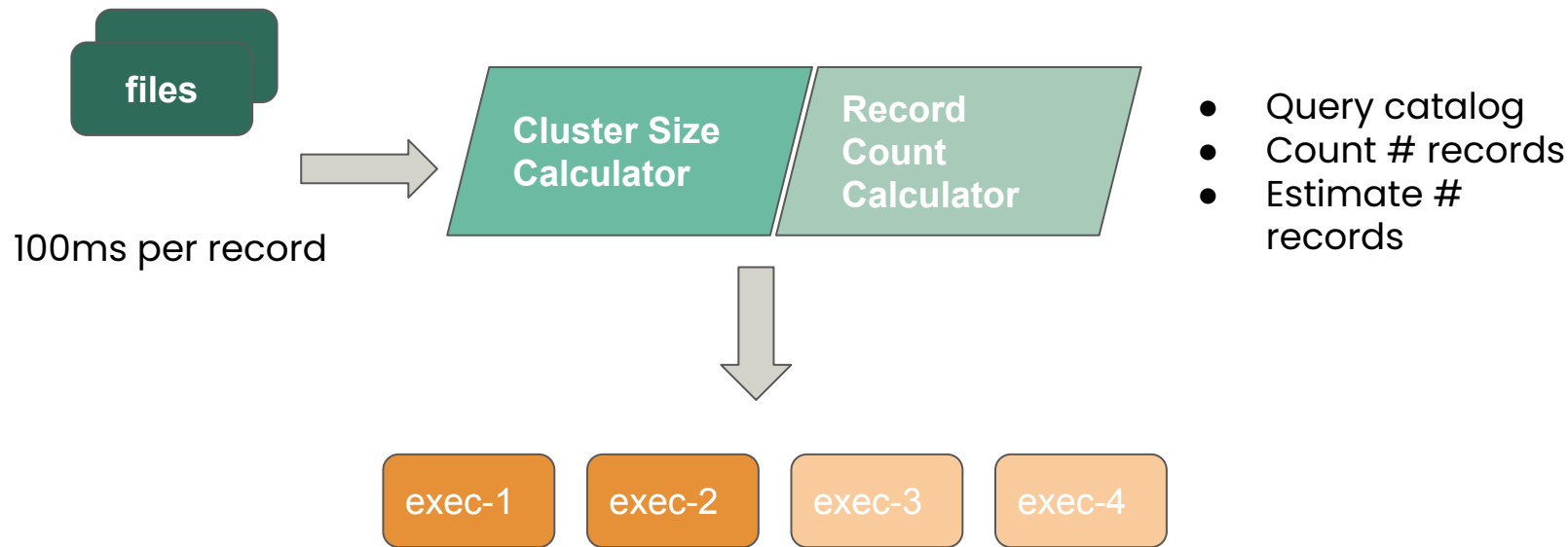
on Argo

- kubernetes-spark-operator
- Installed as k8s custom resource
- Controller orchestrates the creation and management of the Spark job

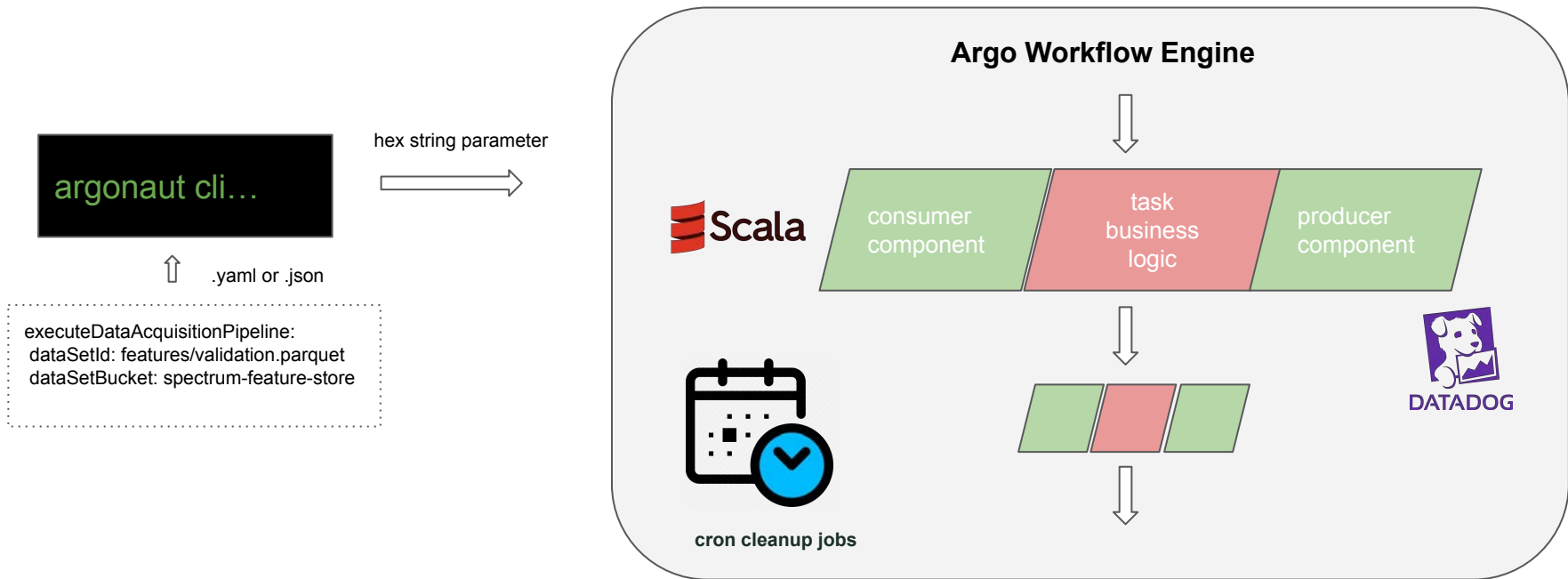
Iroh Demo



Supporting Elastic Scalability



Our Data Pipeline Framework



Argo Events



“Argo Events is an event-driven workflow automation framework for Kubernetes which helps you trigger K8s objects, Argo Workflows, Serverless workloads, etc. on events from a variety of sources like webhooks, S3, schedules, messaging queues, gcp pubsub, sns, sqs, etc.”



Argo Events Demo



Challenges We Encountered

Spark operator is not
native Argo component

yaml definitions are
cumbersome

Newer technology

Minikube setup
challenging but essential

Platform tuning for Spark jobs



Q&A

The image features the letters 'Q&A' in a bold, three-dimensional, metallic font. The characters are rendered with a brushed metal texture and sharp edges, giving them a modern, industrial appearance. They are set against a plain white background, which is itself centered within a larger teal frame. The lighting creates highlights and shadows on the letters, emphasizing their three-dimensional form. On the right side of the teal background, there are faint, overlapping circular lines that add a subtle decorative element to the overall design.