

The background of the image is a vibrant, multi-colored gradient. It features broad diagonal bands of color, including shades of blue, purple, magenta, orange, and yellow, creating a dynamic and modern aesthetic. The AWS re:Invent logo is positioned on the left side of the image, rendered in white text.

AWS  
re:Invent

DEM30-S

# Continuous deployment for ML: The new software development lifecycle

**Diego Oppenheimer**

CEO

Algorithmia

\_\_\_\_\_

The diagram illustrates the distribution of effort between two tasks. On the left, a calendar grid shows the days of the week (S, M, T, W, R, F, S) across several weeks. The first three weeks are entirely purple, representing 'Manual Infrastructure Work'. The last two weeks show a mix of purple and teal, representing 'Model Building & Training'. On the right, two boxes summarize the effort: 'Manual Infrastructure Work' accounts for 75% of the effort, and 'Model Building & Training' accounts for 25% of the effort.

| Day | Manual Infrastructure Work | Model Building & Training |
|-----|----------------------------|---------------------------|
| S   | Yes                        | No                        |
| M   | Yes                        | No                        |
| T   | Yes                        | No                        |
| W   | Yes                        | No                        |
| R   | Yes                        | No                        |
| F   | Yes                        | No                        |
| S   | Yes                        | No                        |

Manual Infrastructure Work 75% effort

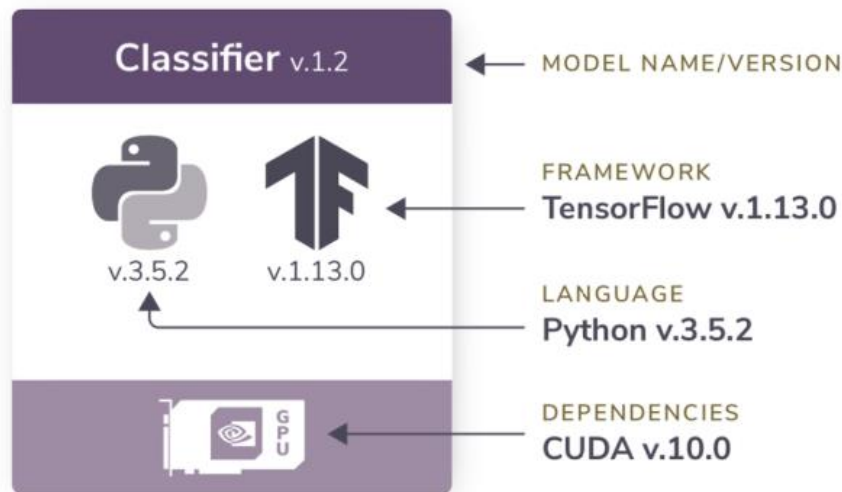
Model Building & Training 25% effort

- **30%:** Supporting different languages and frameworks
- **30%:** Model management tasks such as versioning and reproducibility
- **38%:** Deploying models at necessary scale

Survey of >500 practitioners & managers in summer of 2018

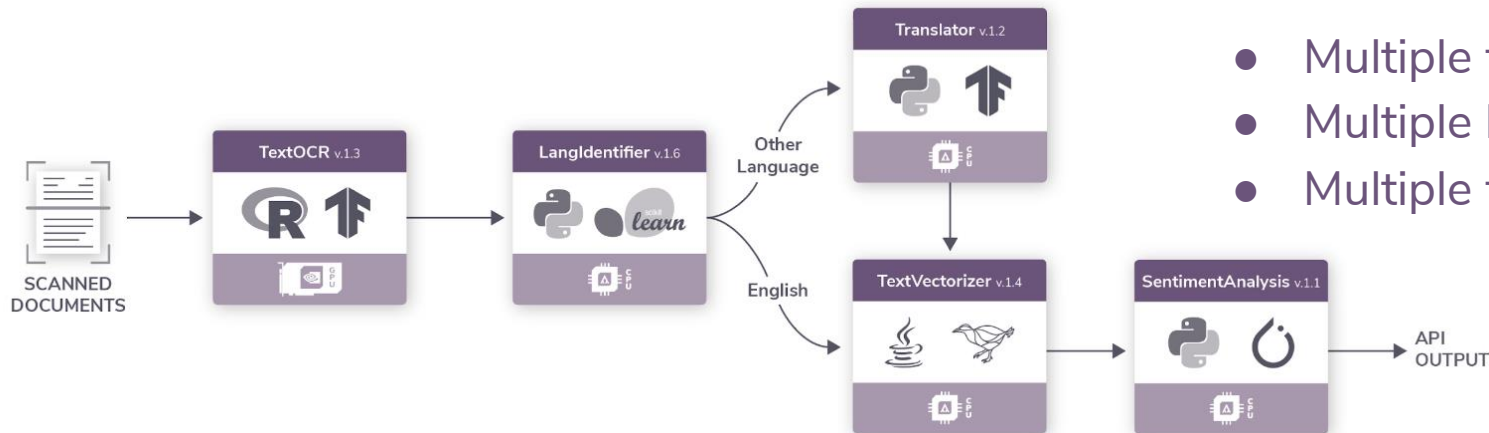
# Heterogeneous tooling & dependencies

---



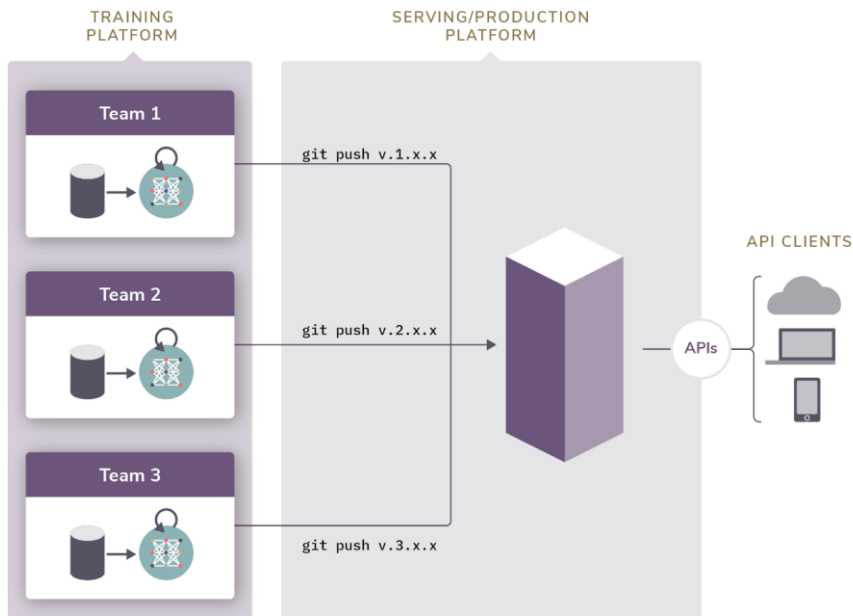
- Dozens of language and framework combinations
- Hardware dependencies (e.g., CUDA) require substantial architecture investment
- New frameworks emerge every year
- Frameworks and languages evolve constantly, requiring ongoing maintenance and testing

# Composability compounds the challenge



- Multiple frameworks
- Multiple languages
- Multiple teams

# Training and production are very different



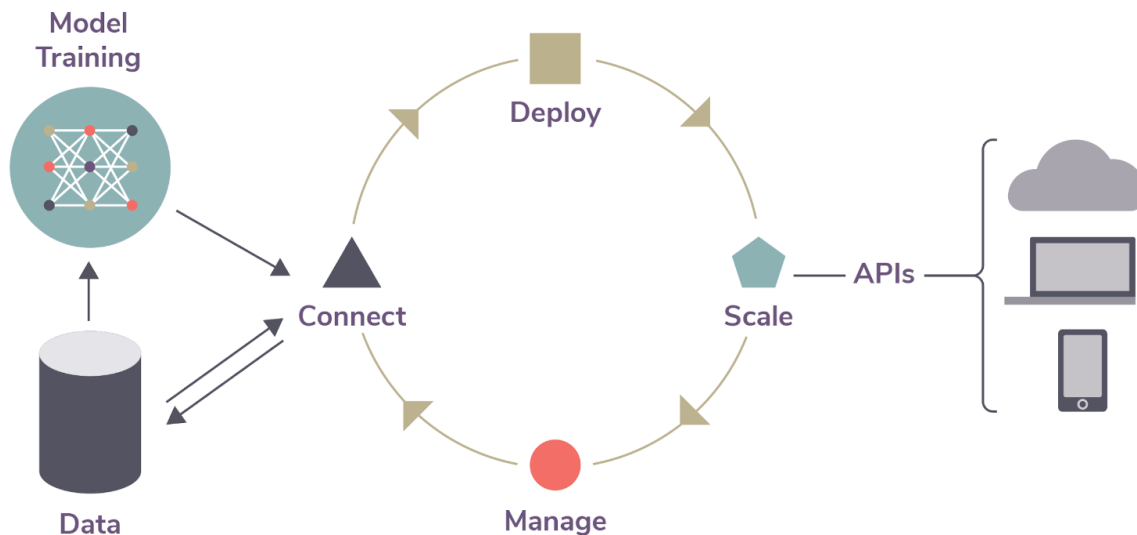
## Training

- Long compute cycle
- Fixed load
- Stateful
- Single user

## Production

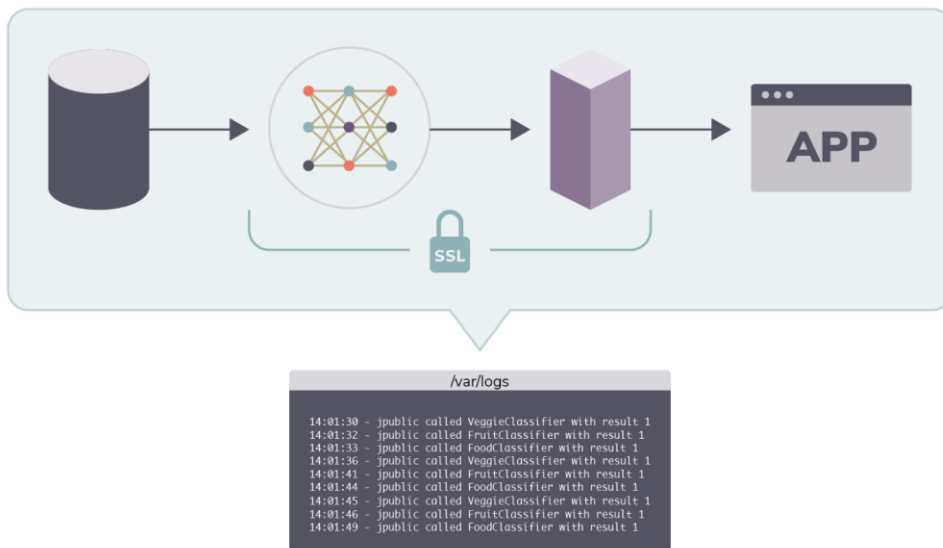
- Short compute bursts
- Elastic
- Stateless
- Many users

# Iteration speed decouples ML from application development



- ML development lifecycle is an evolving ecosystem
- ML moves faster than traditional application development
- ML can introduce breaking changes to applications that consume model output

# Diversity complicates auditability & governance



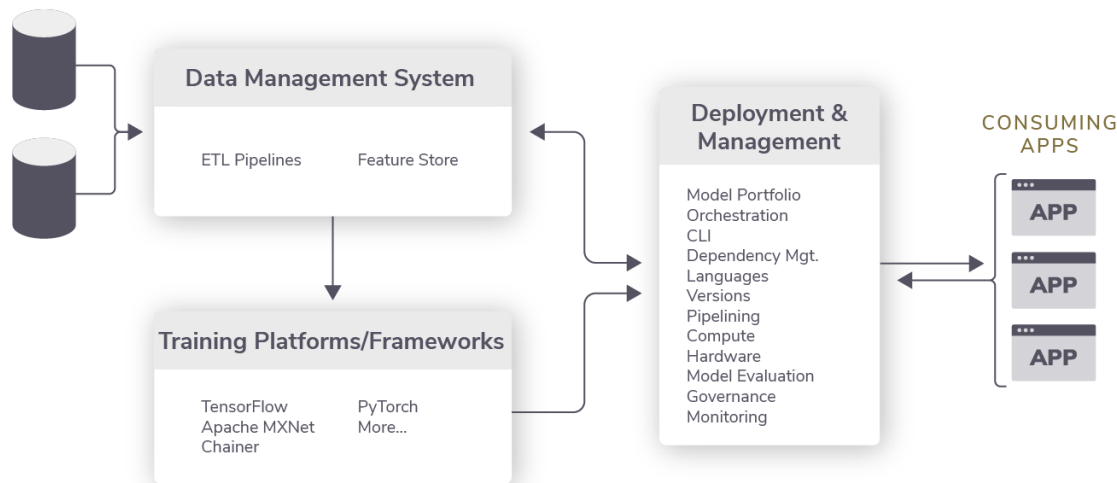
- Internal model usage is difficult to track across multi-model pipelines
- Auditability and access are major security and compliance concerns



# Solution: CD4ML– Automated model deployment

## Programmatically

- **Connect** to data
- **Publish** from training platforms
- **Deploy** models and dependencies
- **Manage** model serving, inference, and compute infrastructure
- **Integrate** with other models and consuming production applications





Demo

```
algorithm_template (master) $ git commit -m "update trained model" digits_classifier.pkl
```

# Q&A

---

Visit us at booth #311

Request a demo at

<https://algorithmia.com/demo>

Download a whitepaper at

<http://bit.ly/AlgoReInvent>

Contact us at

[info@algorithmia.com](mailto:info@algorithmia.com)







# Appendix

```
algorithm_template (master) $ git commit -m "update trained model" digits_classifier.pkl
On branch master
Your branch is ahead of 'origin/master' by 2 commits.
  (use "git push" to publish your local commits)
nothing to commit, working tree clean
algorithm_template (master) $ git push
```

```
[ algorithm_template (master) $ git commit -m "update trained model" digits_classifier.pkl
On branch master
Your branch is ahead of 'origin/master' by 2 commits.
  (use "git push" to publish your local commits)
nothing to commit, working tree clean
[ algorithm_template (master) $ git push
Username for 'https://github.com': peckjon
Password for 'https://peckjon@github.com':
Counting objects: 13, done.
Delta compression using up to 12 threads.
Compressing objects: 100% (12/12), done.
Writing objects: 100% (13/13), 2.05 KiB | 0 bytes/s, done.
Total 13 (delta 3), reused 0 (delta 0)
remote: Resolving deltas: 100% (3/3), completed with 3 local objects.
To https://github.com/algorithmiaio/model-deployment.git
   761323d..ec29803  master -> master
algorithm_template (master) $
```



```

master' by 2 commits.
+ local commits)
clean
t push
: peckjan
hub.com':

threads.
, done.
05 KiB | 0 bytes/s, done.
ta 0)
/3), completed with 3 local objects.
io/model-deployment.git
ster

```

← → ↺

GitHub, Inc. [US] | github.com/algorithmiaio/model-deployment/commit/ec29803855fd5c66f6a...

☆ 🔔 ⚙️

🐱

Search or jump to...

/

Pull requests

Issues

Marketplace

Explore

🔔

+

👤

algorithmiaio / model-deployment

👁️ Unwatch 12

★ Star 3

🍴 Fork 0

<> Code

🔔 Issues 1

🔗 Pull requests 0

▶️ Actions

📁 Projects 0

📖 Wiki

🛡️ Security

📊 Insights

⚙️ Settings

🟡

Merge remote-tracking branch 'origin/master'

master 👤 ec29803

Cancel check suite

Deploy Model on Algorithmia

on: push

● build (3.6)

GitHub Actions / build (3.6) started 20s ago

Search logs

< > ⋮

✔️ Set up job 2s

✔️ Run actions/checkout@v1 3s

✔️ `${{ format('Set up Python {0}', matrix.python-version) }}` 3s

🟡 Deploy model 12s

```
15 Found existing installation: pip 19.2.2
16 Uninstalling pip-19.2.2:
17 Successfully uninstalled pip-19.2.2
18 Successfully installed pip-19.2.3
19 CREATING data:///***digit_recognition
20 UPLOADING model to data:///***digit_recognition/digits_classifier.pkl
21 CREATING ***digit_recognition_20190917203726
22 {'details': {'label': 'Digit Recognition',
23             'summary': 'Demo only: Provide an image URL as input, and this '
24                       'model will attempt to identify the number shown.',
25             'tagline': 'Simple digit recognition model, for demo purposes '
26                       'only'},
27  'name': 'digit_recognition_20190917203726',
28  'self_link': None,
```

```

aster' by 2 commits.
, local commits)
lean
t push
: peckjon
hub.com':

threads.
, done.
05 KiB | 0 bytes/s, done.
ta 0)
/3), completed with 3 local objects.
io/model-deployment.git
ster

```

GitHub, Inc. [US] | github.com/algorithmiaio/model-deployment/commit/ec29803855fd5c66f6a... ☆

Search or jump to... Pull requests Issues Marketplace Explore

algorithmiaio / model-deployment

Unwatch 12 Star 3 Fork 0

Code Issues 1 Pull requests 0 Actions Projects 0 Wiki Security Insights Settings

✓ Merge remote-tracking branch 'origin/master'

Deploy Model on Algorithmia

on: push

✓ build (3.6)

GitHub Actions / build (3.6) successful 1 minute ago in 51s

Search logs

< > ...

✓ Set up job2s

✓ Run actions/checkout@v13s

✓ `${{ format('Set up Python {0}', matrix.python-version) }}`3s

✓ Deploy model43s

✓ Complete job0s

1 Cleaning up orphan processes

```
aster' by 2 commits.  
local commits)  
lean  
t push  
: peckjon  
hub.com':  
  
threads.  
, done.  
05 KiB | 0 bytes/s, done.  
ta 0)  
/3), completed with 3 local objects.  
io/model-deployment.git  
ster
```



jpeck / digit\_recognition\_20190917203726 /  
0.1.0

Star: 0

Unfollow: 1

Overview

Docs

Discussion

Listeners

Builds

Settings

Source Code →

## Digit Recognition [Royalty Free](#) [Edit](#)

Simple digit recognition model, for demo purposes only [Edit](#)

No Tags [Edit](#)

LANGUAGE

Python 3.x

METRICS

Avg call duration - N/A

PERMISSIONS

[Private](#)

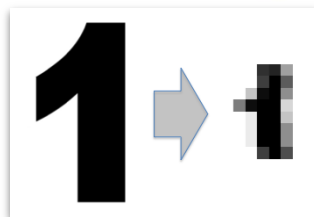
[Algorithmia Platform License](#)

[Internet Access](#)

[Pipelining](#)



The first image or video in your README.md will appear here  
(editable under the [Docs](#) tab)



## Run an Example

INPUT

```
"https://github.com/algorithmiaio/sample-  
apps/blob/master/algo-dev-  
demo/digit_recognition/images/digit_1.png?  
raw=true"
```

[Copy](#)

OUTPUT

1

[Copy](#)

```
aster' by 2 commits.  
local commits)  
lean  
t push  
: peckjon  
hub.com':  
  
threads.  
, done.  
05 KiB | 0 bytes/s, done.  
ta 0)  
, completed with 3 local objects.  
io/model-deployment.git  
ster
```



# Digit Recognition Royalty Free Edit

Simple digit recognition model, for demo purposes only Edit

No Tags Edit

## LANGUAGE

Python 3.x

## METRICS

Avg call duration - N/A

## PERMISSIONS

Private Algorithmia Platform License Internet Access Pipelining

## Run an Example

### INPUT

```
"https://github.com/algorithmiaio/sample-  
apps/blob/master/algo-dev-  
demo/digit_recognition/images/digit_1.png?  
raw=true"
```

Copy

### OUTPUT

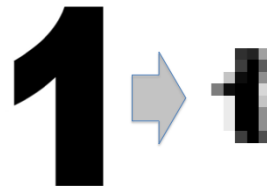
```
** Running algorithm...
```

Copy

Save As Default

Run Example

The first image or video in your README.md will appear here  
(editable under the Docs tab)



```
aster' by 2 commits.  
local commits)  
lean  
t push  
: peckjon  
hub.com':  
  
threads.  
, done.  
05 KiB | 0 bytes/s, done.  
ta 0)  
/3), completed with 3 local objects.  
io/model-deployment.git  
ster
```



# Digit Recognition Royalty Free Edit

Simple digit recognition model, for demo purposes only Edit

No Tags Edit

## LANGUAGE

Python 3.x

## METRICS

Avg call duration - N/A

## PERMISSIONS

Private Algorithmia Platform License Internet Access Pipelining

## Run an Example

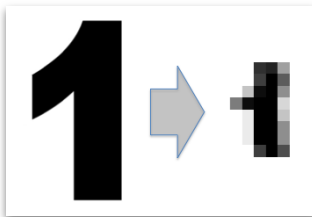
### INPUT

```
"https://github.com/algorithmiaio/sample-  
apps/blob/master/algo-dev-  
demo/digit_recognition/images/digit_1.png?  
raw=true"
```

Copy

Save As Default

Run Example



### OUTPUT

1

Copy



Please complete the session  
survey in the mobile app.