

Are you ready for **MLOps**?



Jetze Schuurmans
Jeroen Overschie

PyData London — 16th of June 2024



Xebia | Data

Jeroen Overschie
Machine Learning Engineer



Xebia | Data

Jetze Schuurmans
Machine Learning Engineer

What does
it mean to
be **ready**?

To find out,
let's first look
at what **MLOps**
promised us.



The promise of MLOps

Why MLOps?

The promise of MLOps

Why MLOps?

Because many companies deal with common **challenges**.



Long development cycles

To create models, update models or add features



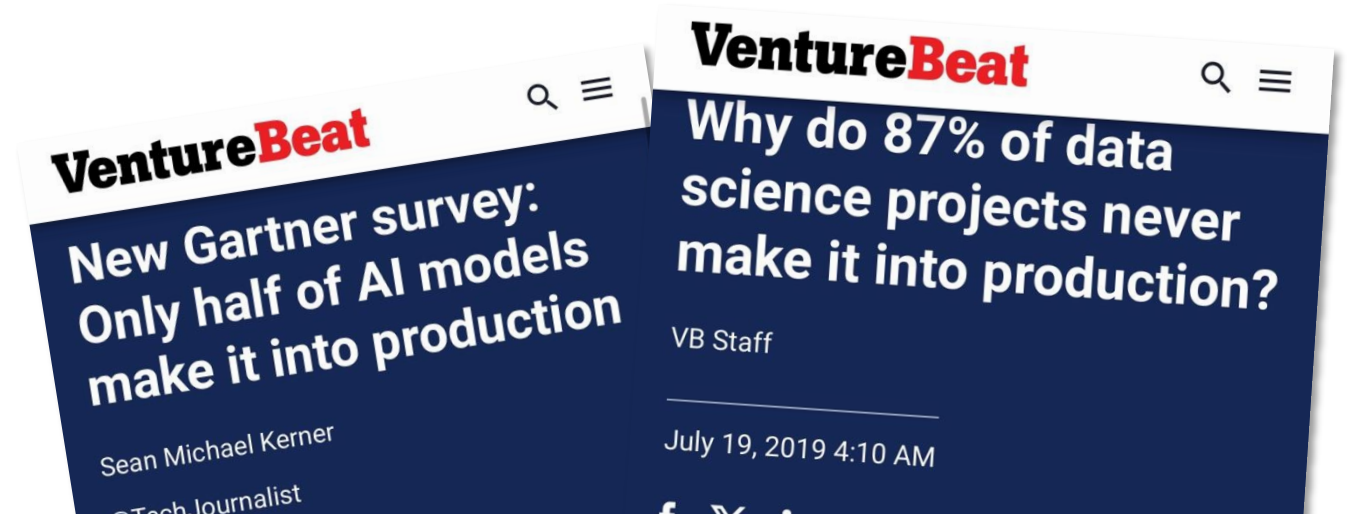
Lacking feedback

On model performance and added value



Few models in production

Unreliable solutions



The promise of MLOps

Why MLOps?

Because many companies deal with common **challenges**.



Long development cycles

To create models, update models or add features



Lacking feedback

On model performance and added value

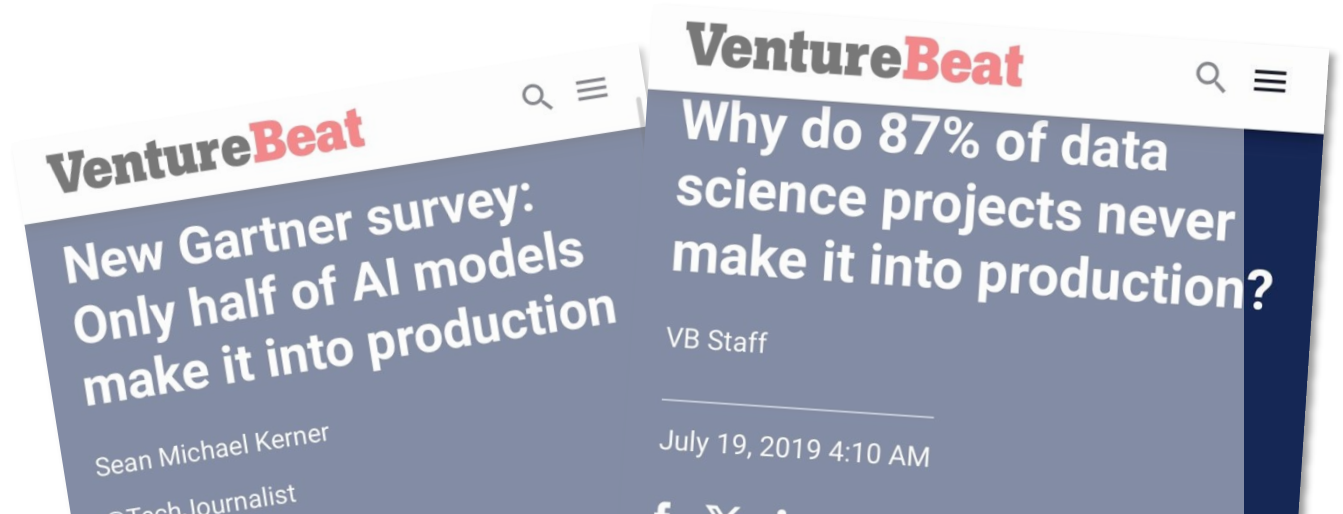


How come?

Why is this still difficult these days?

Few models in production

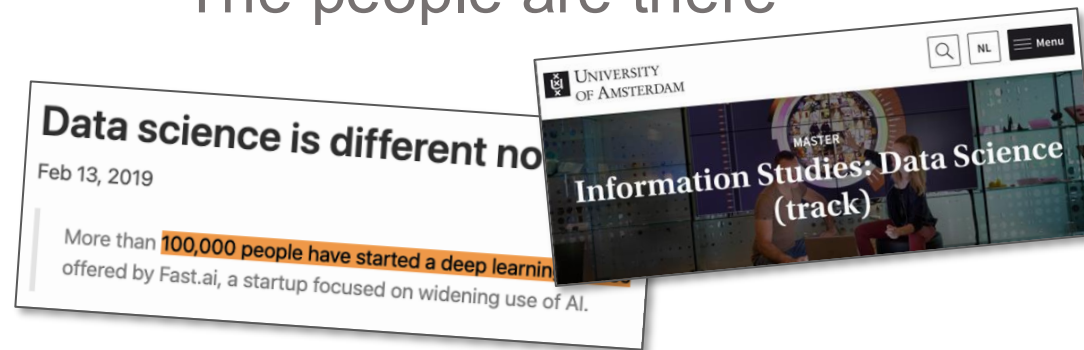
Unreliable solutions



The tech is there



The people are there



How come?

Why is this still difficult these days?

The meetups
are there 😊

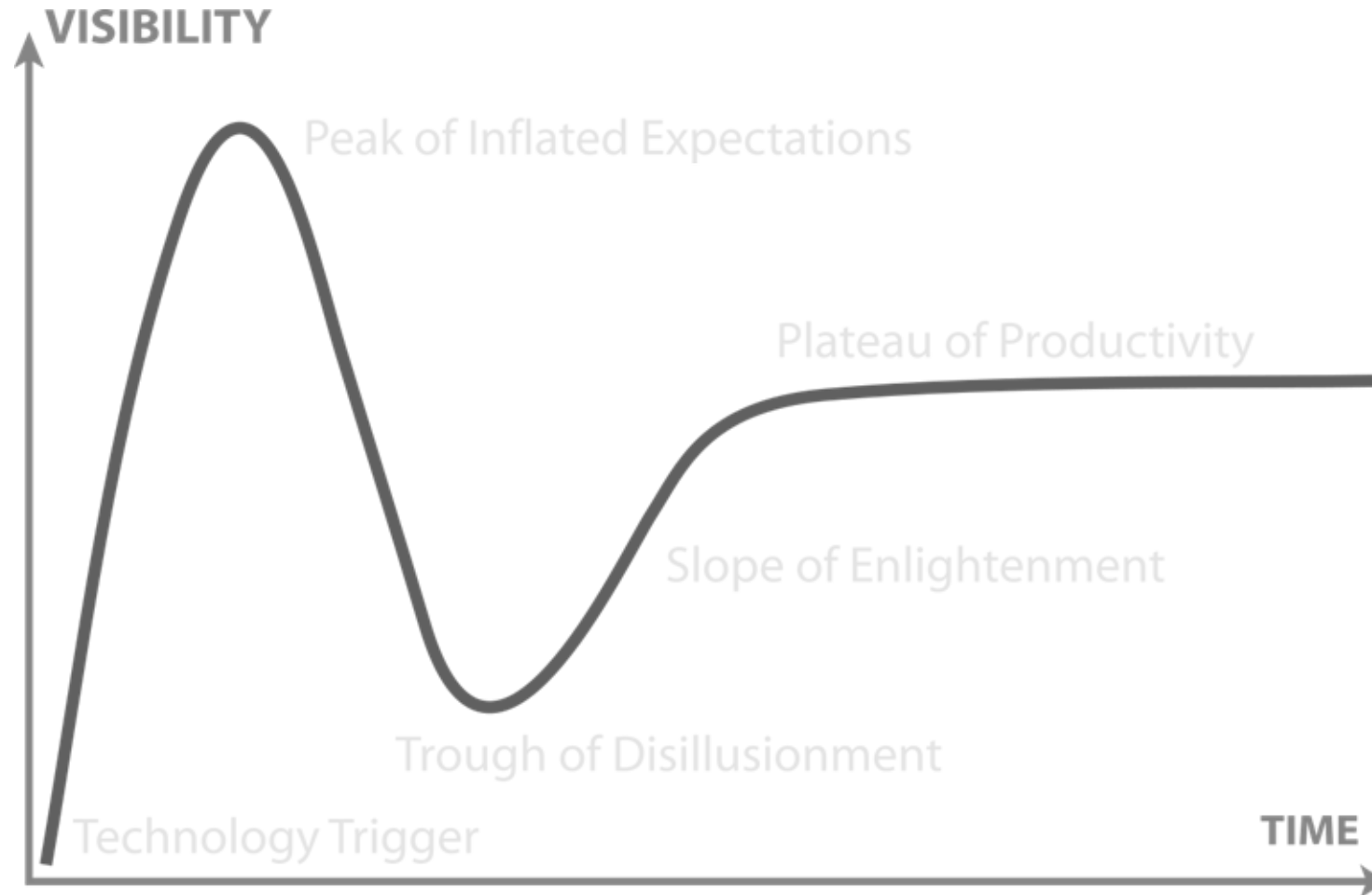


The promise of MLOps



The promise of **MLOps**

Hype cycle



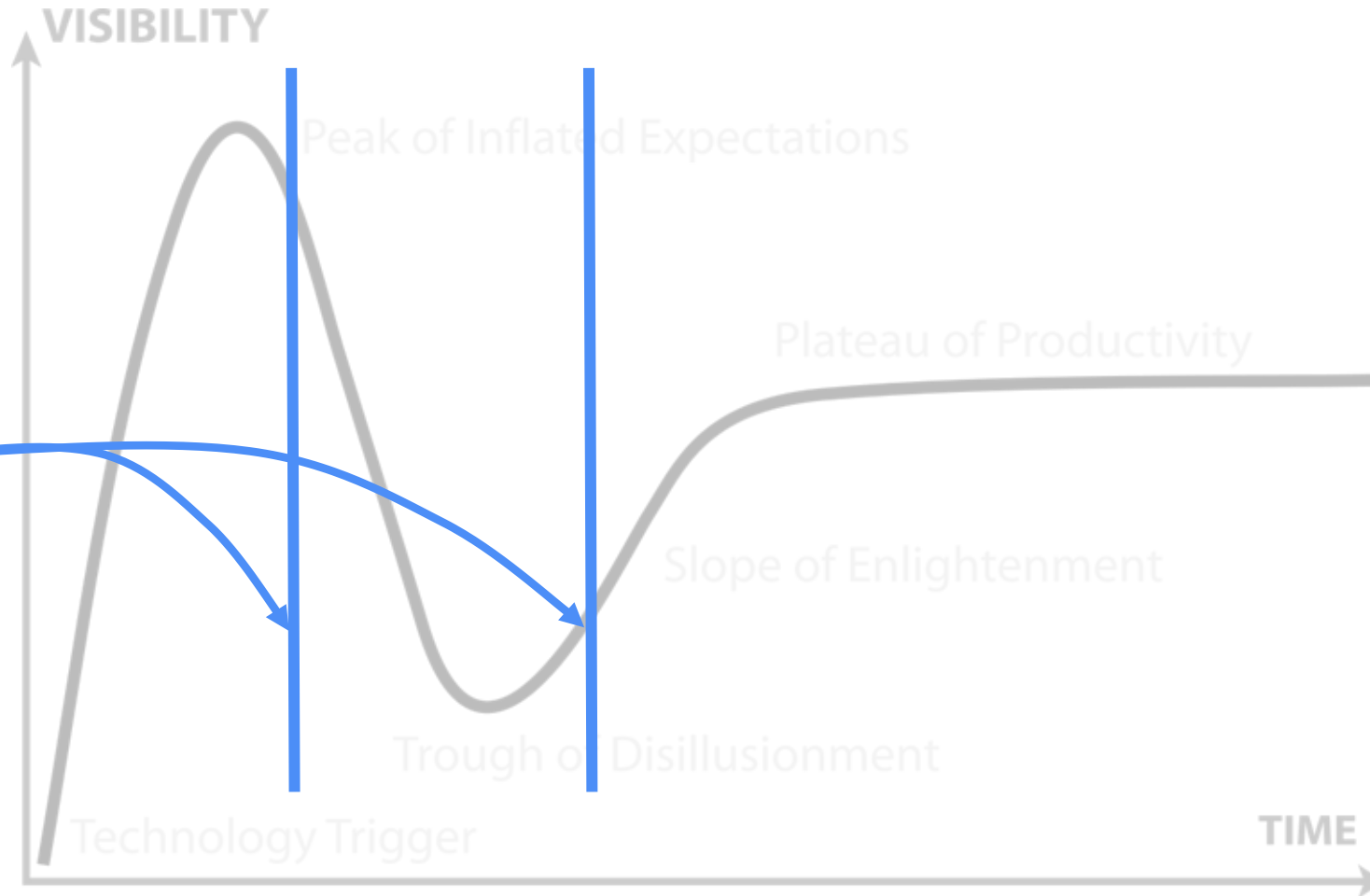
About the hype cycle
The *Gartner* hype cycle is a graphical presentation developed, used and branded by the American research, advisory and information technology firm Gartner to represent **the maturity, adoption, and social application of specific technologies**. The hype cycle claims to provide a graphical and conceptual presentation of the maturity of emerging technologies through five phases. ^[1]

[1] https://en.wikipedia.org/wiki/Gartner_hype_cycle

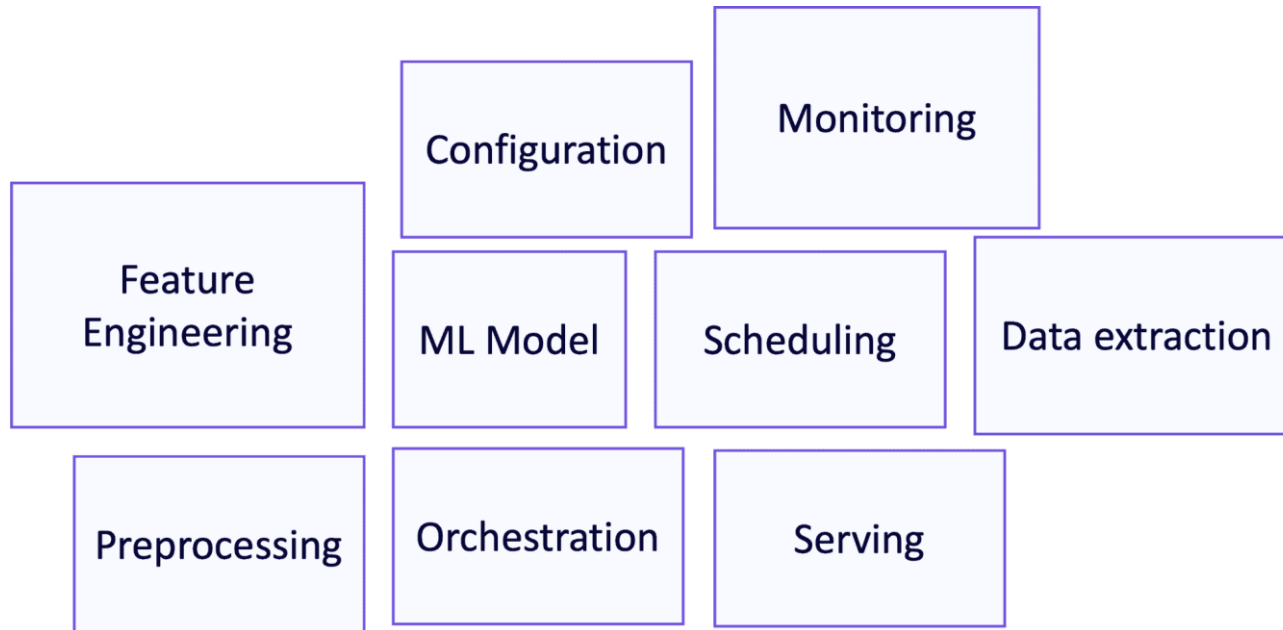
The promise of MLOps

Hype cycle

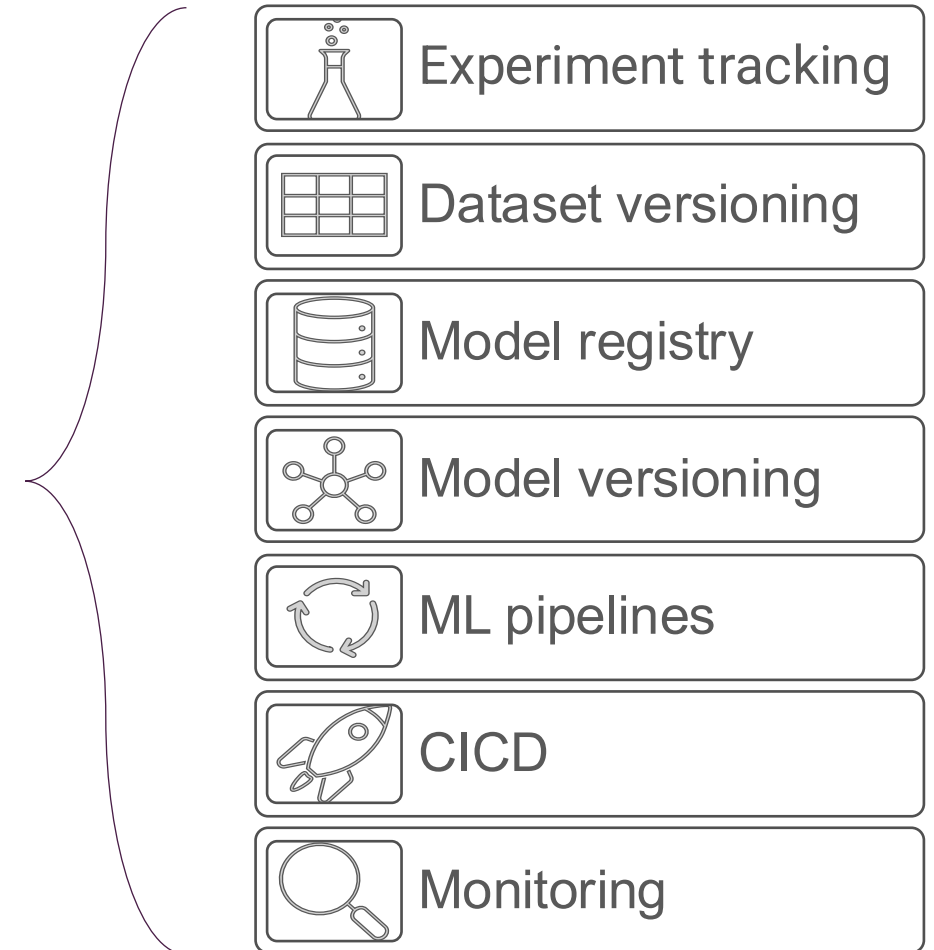
MLOps



The promise of MLOps



<https://xebia.com/blog/mlops-why-and-how-to-build-end-to-end-product-teams/>

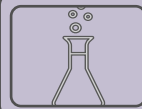
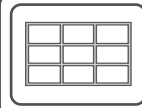
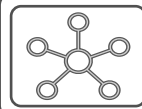
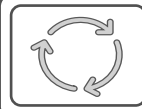
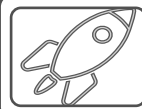


The promise of MLOps

The screenshot displays the mlflow web interface. On the left, there's a sidebar with 'Experiments' and 'Models' tabs. The main area shows the 'mlflow-demo' experiment details, including a description, filters, and a table of runs. The table lists runs with columns for Start Time, Duration, Run Name, User, Source, Version, Models, Metrics (accuracy), and Parameters (depth).

	Start Time	Duration	Run Name	User	Source	Version	Models	Metrics	Parameters
☐	2 minutes ago	9.4s	run_4	KayJan	C:\Program	-	sklearn	1	20
☐	2 minutes ago	11.9s	run_3	KayJan	C:\Program	-	sklearn	1	10
☐	2 minutes ago	12.4s	run_2	KayJan	C:\Program	-	sklearn	0.967	5
☐	3 minutes ago	16.8s	run_1	KayJan	C:\Program	-	sklearn	0.967	2
☐	3 minutes ago	10.7s	run_0	KayJan	C:\Program	-	sklearn	0.633	1

<https://towardsdatascience.com/experiment-tracking-with-mlflow-in-10-minutes-f7c2128b8f2c>

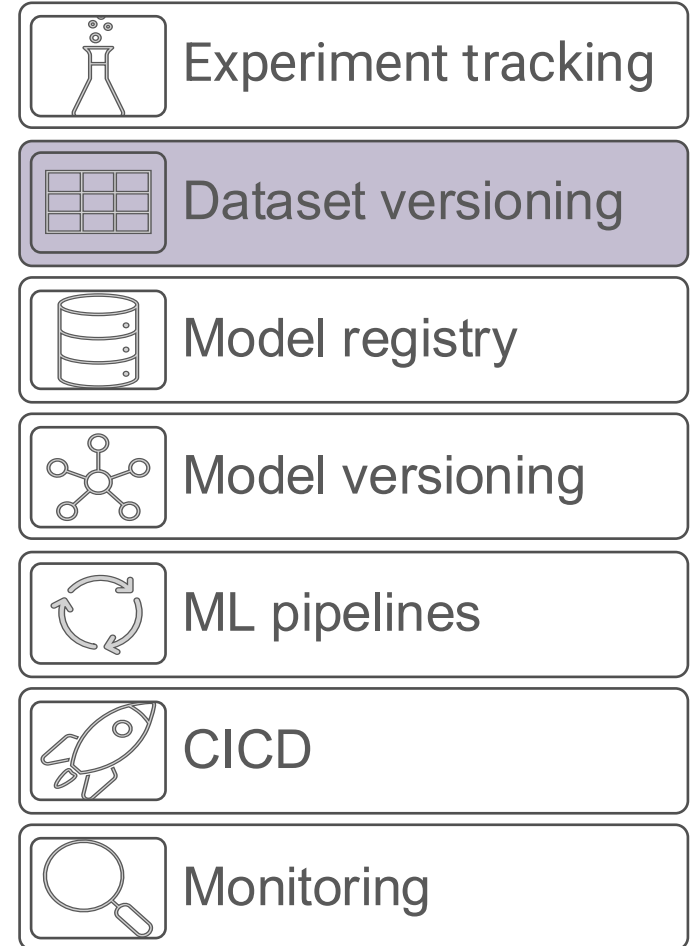
-  Experiment tracking
-  Dataset versioning
-  Model registry
-  Model versioning
-  ML pipelines
-  CI/CD
-  Monitoring

The promise of MLOps

The screenshot displays the Azure Machine Learning workspace interface. The left sidebar contains navigation options: New, Home, Author, Notebooks, Automated ML, Designer, Assets, Datasets, Experiments (highlighted), Pipelines, Models, Endpoints, Manage, Compute, and Environments. The main area shows the breadcrumb path: mayworkspace > Experiments > keras-mnist > keras-mnist_1570739212_92bc7afd. Below this, it indicates 'Run 24' with a 'Switch to old experience' link. A 'Refresh' button is present. The 'Properties' tab is active, showing details for 'Run 24'. The 'Input datasets' section is highlighted with a red box, showing 'keras-mnist'. The 'Metrics' tab shows accuracy and loss values.

Properties	Metrics
Status Completed	Accuracy Vector
Created Oct 10, 2019 1:31 PM	Final test accuracy 0.9794999957084656
Duration 1m 55.69s	Final test loss 0.17366168976166854
Compute target local	Loss Vector
Run ID keras-mnist_1570739212_92bc7afd	Loss v.s. Accuracy Image
Run number 24	
Script name keras_mnist.py	
Input datasets keras-mnist	

<https://learn.microsoft.com/en-us/azure/machine-learning/how-to-version-track-datasets?view=azureml-api-1>



The promise of MLOps

Vertex AI

TOOLS

Dashboard

Workbench

Pipelines

DATA

Feature Store

Datasets

Labeling tasks

MODEL DEVELOPMENT

Training

Experiments

Metadata

DEPLOY AND USE

Endpoints

Model Registry

Batch predictions

Matching Engine

Model Registry

CREATE

IMPORT

Models are built from your datasets or unmanaged data sources. There are many different types of machine learning models available on Vertex AI, depending on your use case and level of experience with machine learning. [Learn more](#)

Region

us-central1 (Iowa)

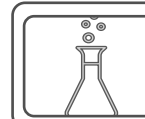
Filter

Deployment status: Deployed on Vertex AI

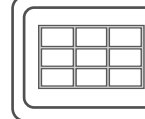
Enter a property name

Name	Deployment status	Description	Default version	Versions	Type
abalone-model	Deployed on Vertex AI	—	1	1	Custom trained
automl_object_detection_quickstart (2)_2020924155613	Deployed on Vertex AI	—	1	1	Object detection
automlUntitled_15997825_20200910050743	Deployed on Vertex AI	—	1	1	Tabular
automl_with_weight_column_20200916_peris (7)	Deployed on Vertex AI	—	1	1	Tabular
automl_ybo_427_automl_1	Deployed on Vertex AI	—	1	1	Object detection
automl-tabular-model-upload-4866906856490008576-3750702443181113344	Deployed on Vertex AI	—	1	1	Imported
automl-tabular-model-upload-6026042800817045504-2775550528308903936	Deployed on Vertex AI	—	1	1	Imported
automl-tabular-model-upload-6099367032250171392-1464271861511618560	Deployed on Vertex AI	—	1	1	Imported
credit_card_default_1619536651363_2021524173021_limit_balance_regression	Deployed on Vertex AI	—	1	1	Tabular
credit_card_default_1619536651363_limit_balance_regression	Deployed on Vertex AI	—	1	1	Tabular
custom_batch_prediction_test_model_20210303	Deployed on Vertex AI	—	1	1	Imported
custom_job_with_gpus_0610	Deployed on Vertex AI	—	1	1	Custom trained
custom_model_tf_flowers	Deployed on Vertex AI	—	1	1	Imported
custom_model_tf_gpu_flowers	Deployed on Vertex AI	—	1	1	Imported
flowers_2021119233932	Deployed on Vertex AI	—	1	1	Image classification
flowers_2021119234048_edge	Deployed on Vertex AI	—	1	1	Image classification

<https://cloud.google.com/blog/products/ai-machine-learning/vertex-ai-model-registry>



Experiment tracking



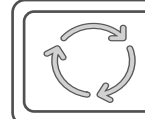
Dataset versioning



Model registry



Model versioning



ML pipelines



CICD



Monitoring

The promise of MLOps

Vertex AI

Churn classifier model [EDIT DETAILS](#)

Model description: A churn prediction model for real-time inference

Region: us-central1 (Iowa)

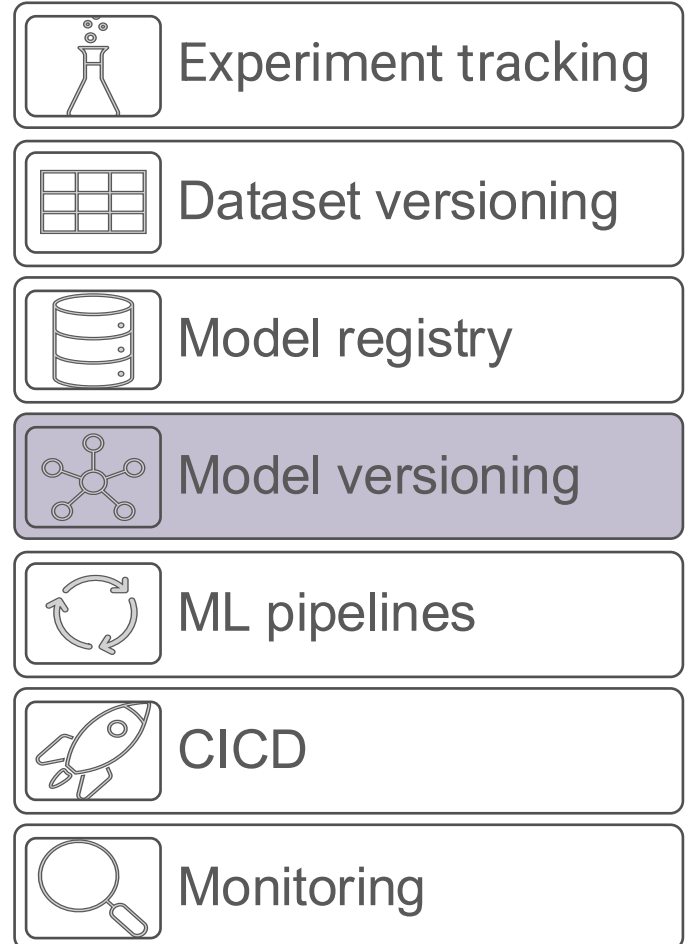
Model labels: —

Versions

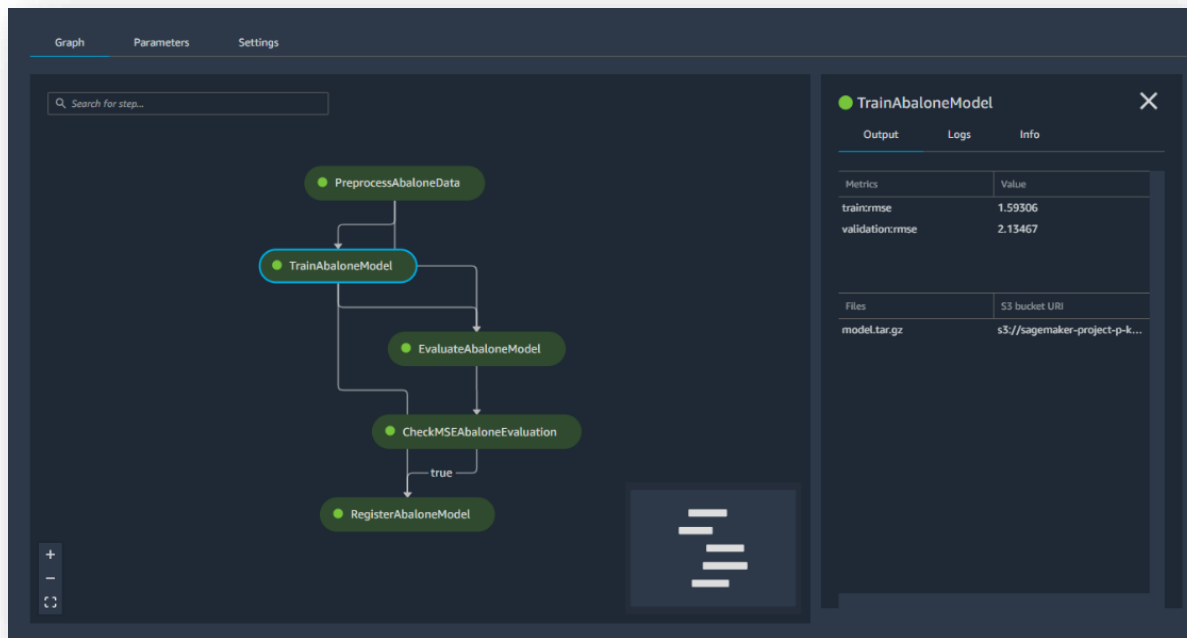
Filter: Enter a property name

Version ID	Alias	Status	Description	Endpoints	Created	Labels
3	rfor experimental	Ready	Random Forest	—	Oct 5, 2022, 5:45:10 PM	created_by: data_scienc... team: data_office
2	gbtree staging	Ready	Gradient Boosting	—	Oct 5, 2022, 5:35:06 PM	created_by: data_scienc... team: data_office
1	default logistic_reg production	Ready	Logistic Regression	—	Oct 5, 2022, 5:33:58 PM	created_by: data_scienc... team: data_office

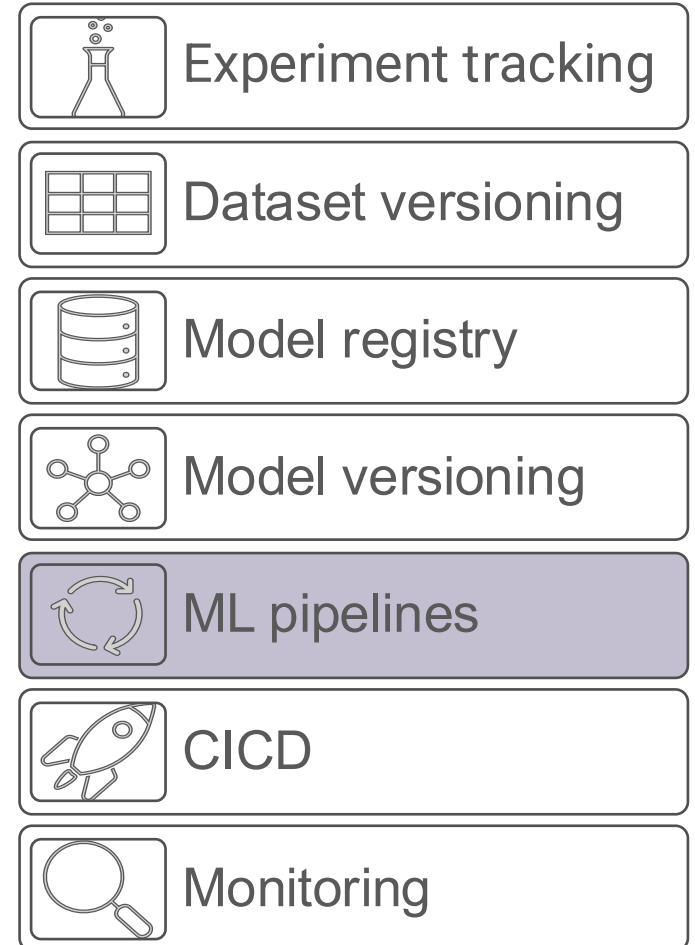
<https://cloud.google.com/blog/products/ai-machine-learning/vertex-ai-model-registry>



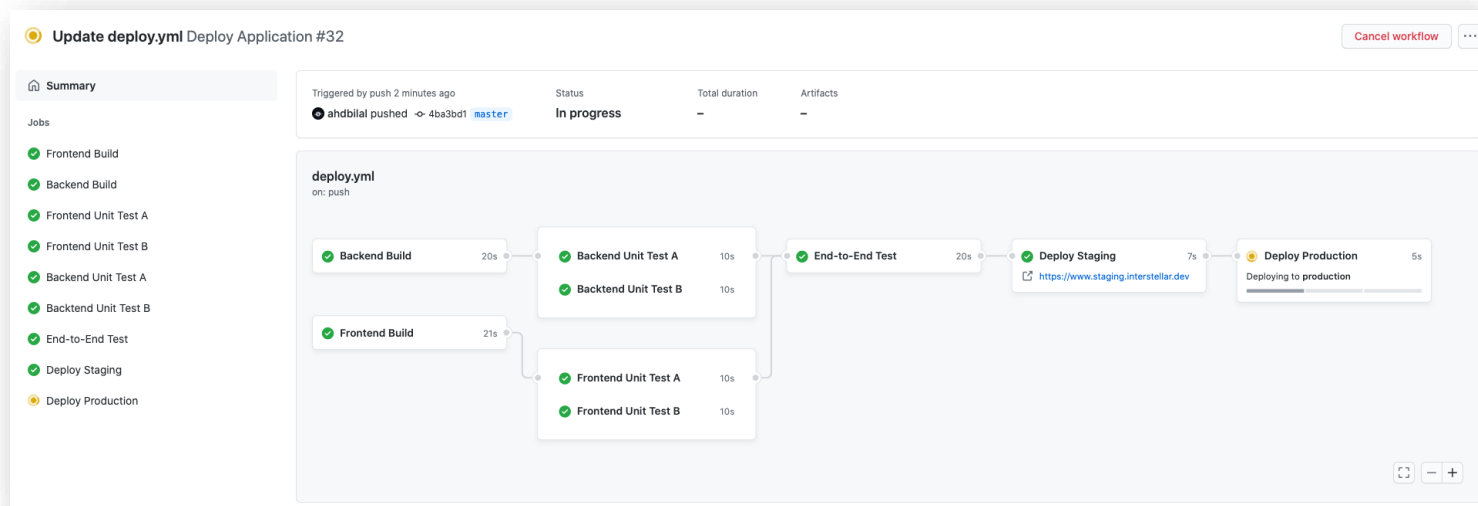
The promise of MLOps



<https://www.amazonaws.cn/en/sagemaker/pipelines/>



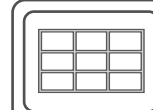
The promise of MLOps



<https://github.blog/change-log/2020-12-08-github-actions-workflow-visualization/>



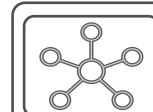
Experiment tracking



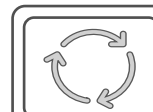
Dataset versioning



Model registry



Model versioning



ML pipelines

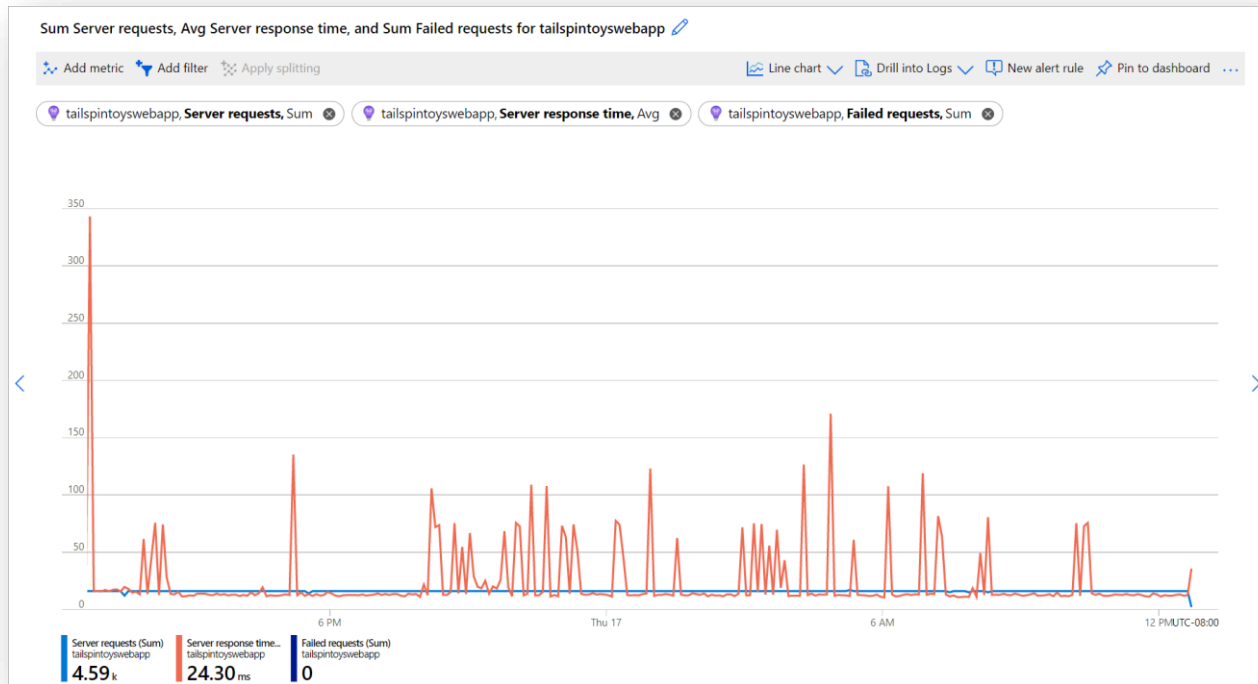


CI/CD

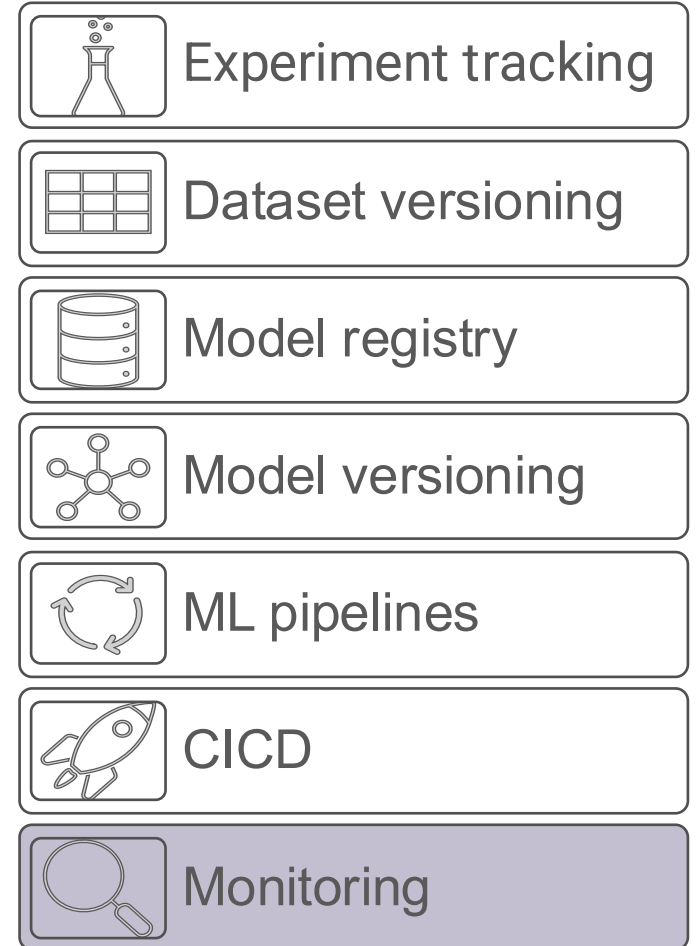


Monitoring

The promise of MLOps



<https://learn.microsoft.com/en-us/azure/azure-monitor/essentials/data-platform-metrics>



The promise of MLOps

Cloud providers offer
solutions for MLOps.



Azure Machine Learning



Vertex AI



Amazon SageMaker



databricks

The promise of MLOps

Why MLOps?

Because many companies deal with common **challenges**.



Long development cycles

To create models, update



Lacking feedback

On model performance



Few models in production

Unreliable

That sounds great but,
Where's this all come from?

MLOps promises



Speed-up development with automation and collaboration



Focus on rapid feedback and iterative development

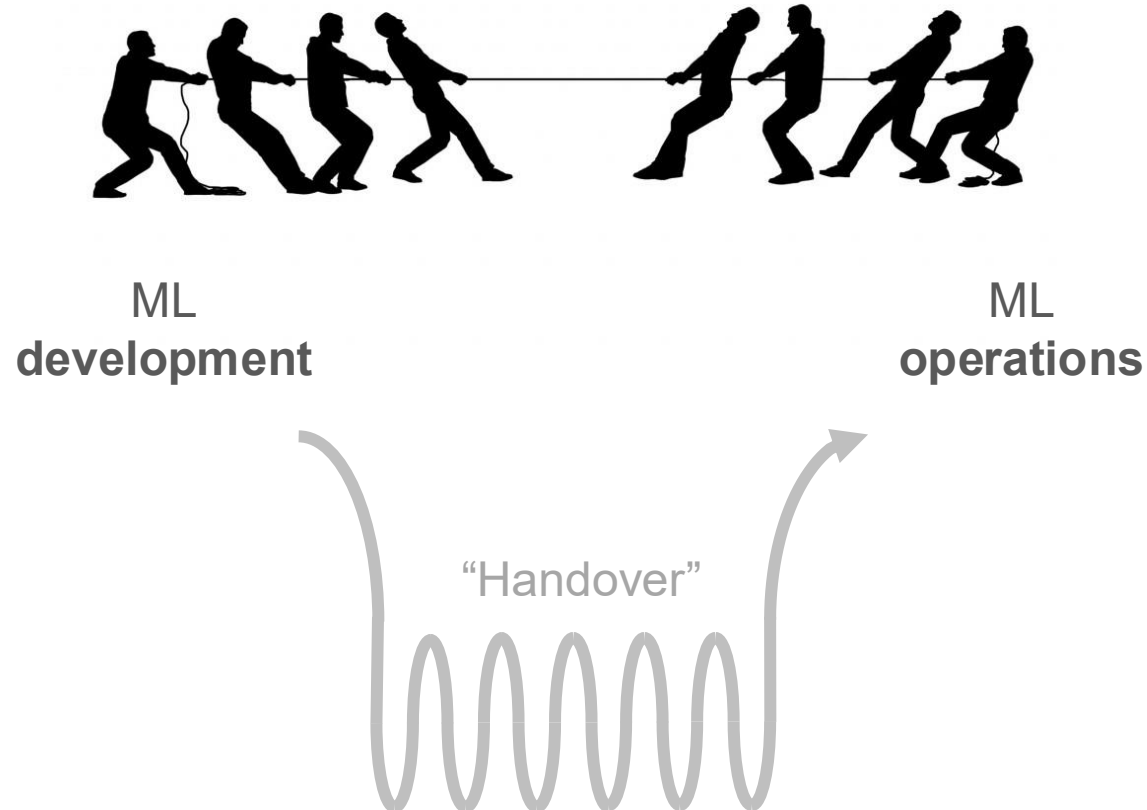


Enable autonomy for end-to-end product teams

The promise of **MLOps**

That all sounds very promising,
but where does this come from?
What are the biggest challenges?

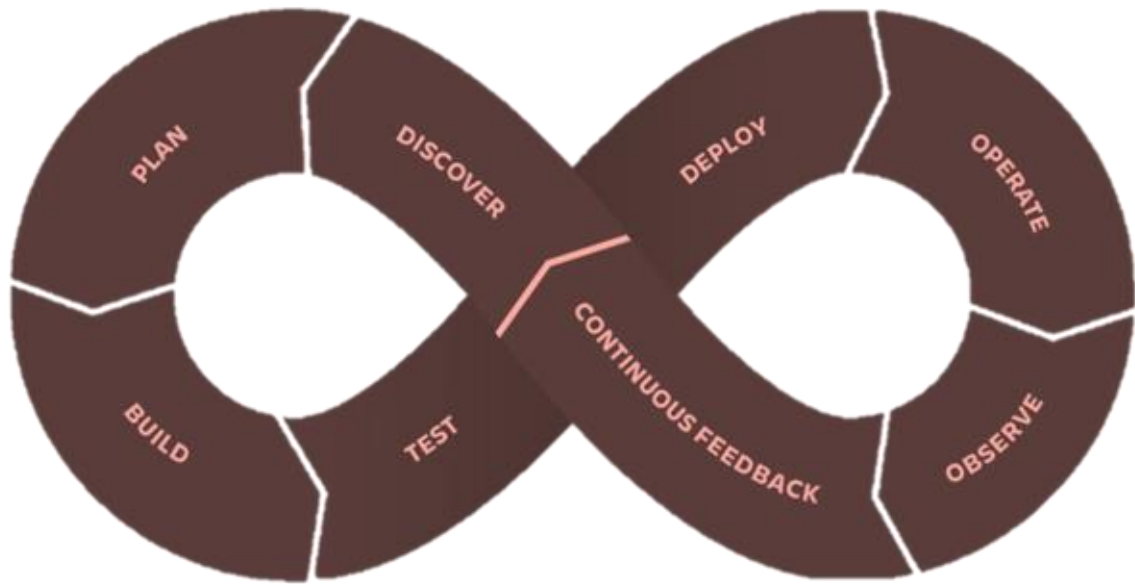
There's **tension**
between
development
and **operations**



A thing called DevOps

Since 2007

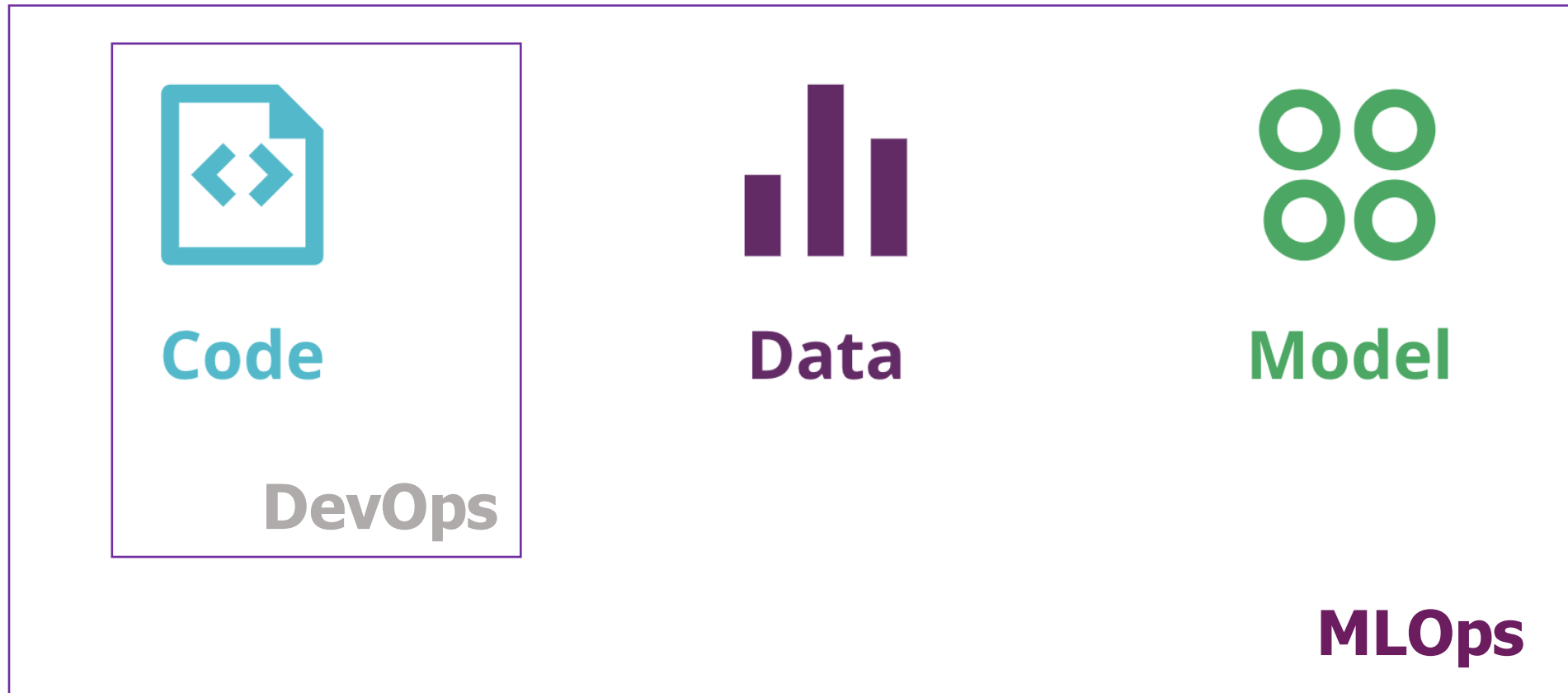
Bring together **Dev** and **Ops**



5 key principles:

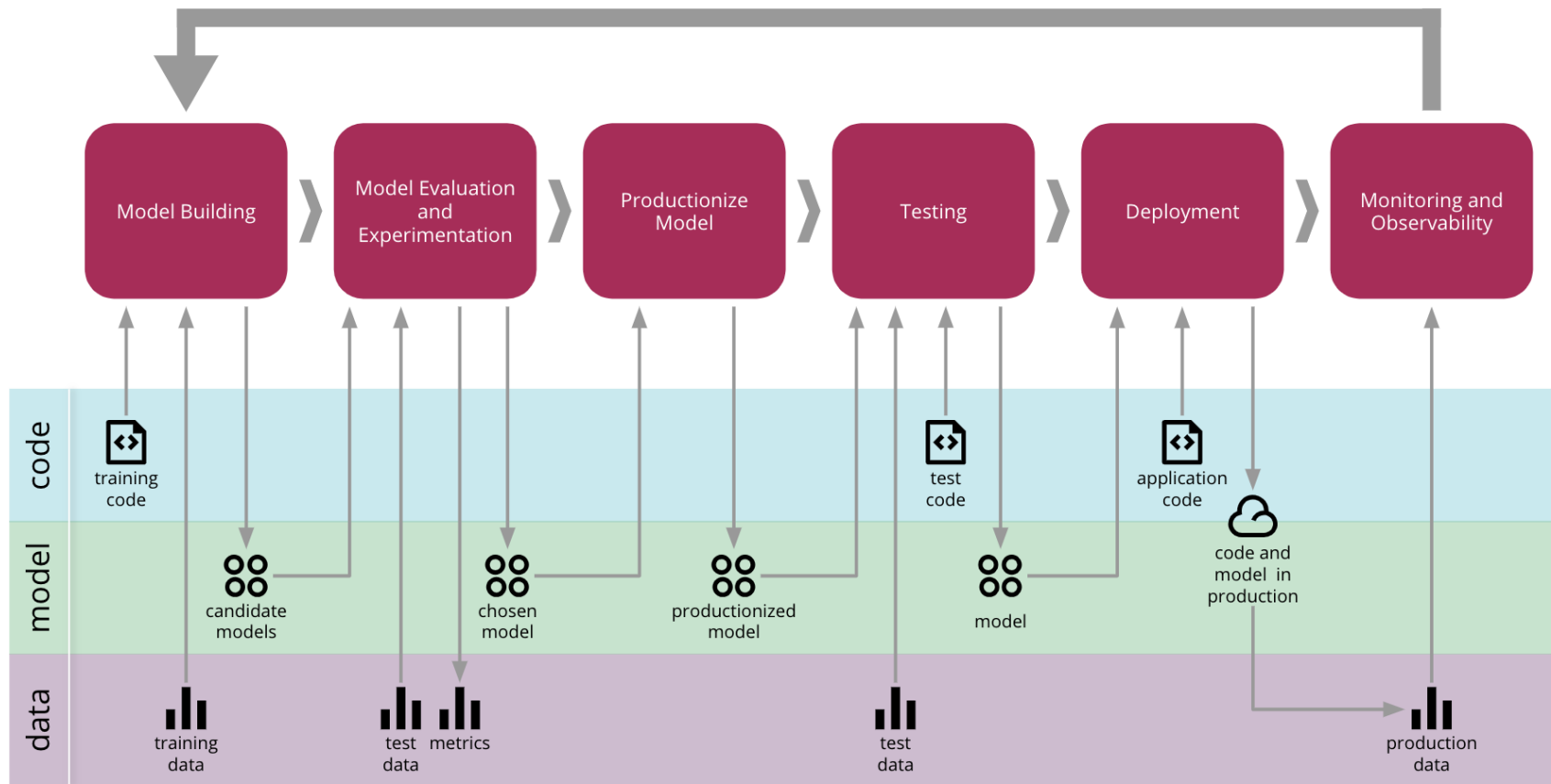
1. Collaborate 🤝
2. Automate ⚙️
3. Continuous Improvements 📈
4. Customer-centric action 👑
5. Create with the end in mind 🚩

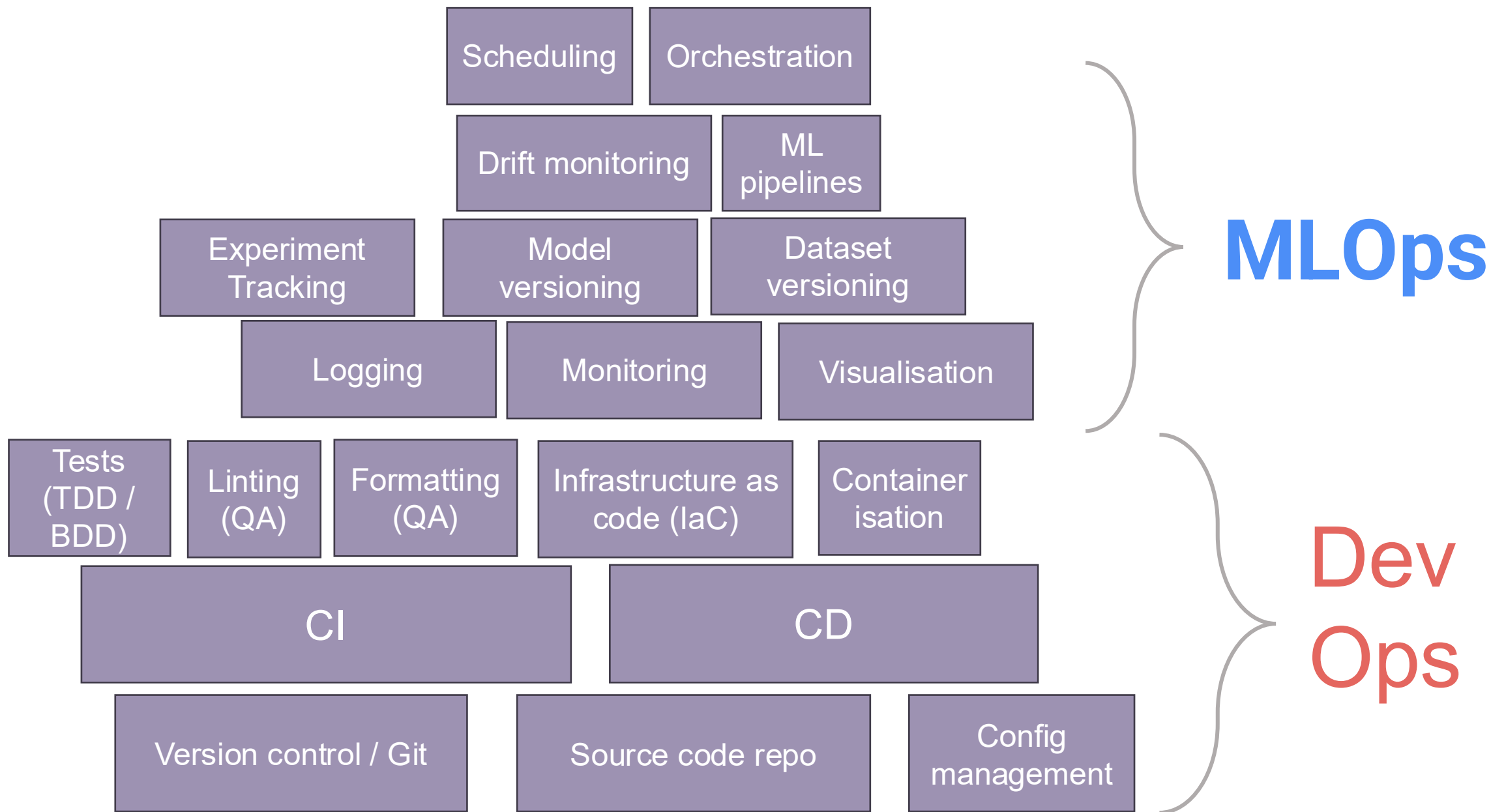
DevOps vs. MLOps



DevOps & MLOps

You need both in a Machine Learning Pipeline







MLOps

**Dev
Ops**

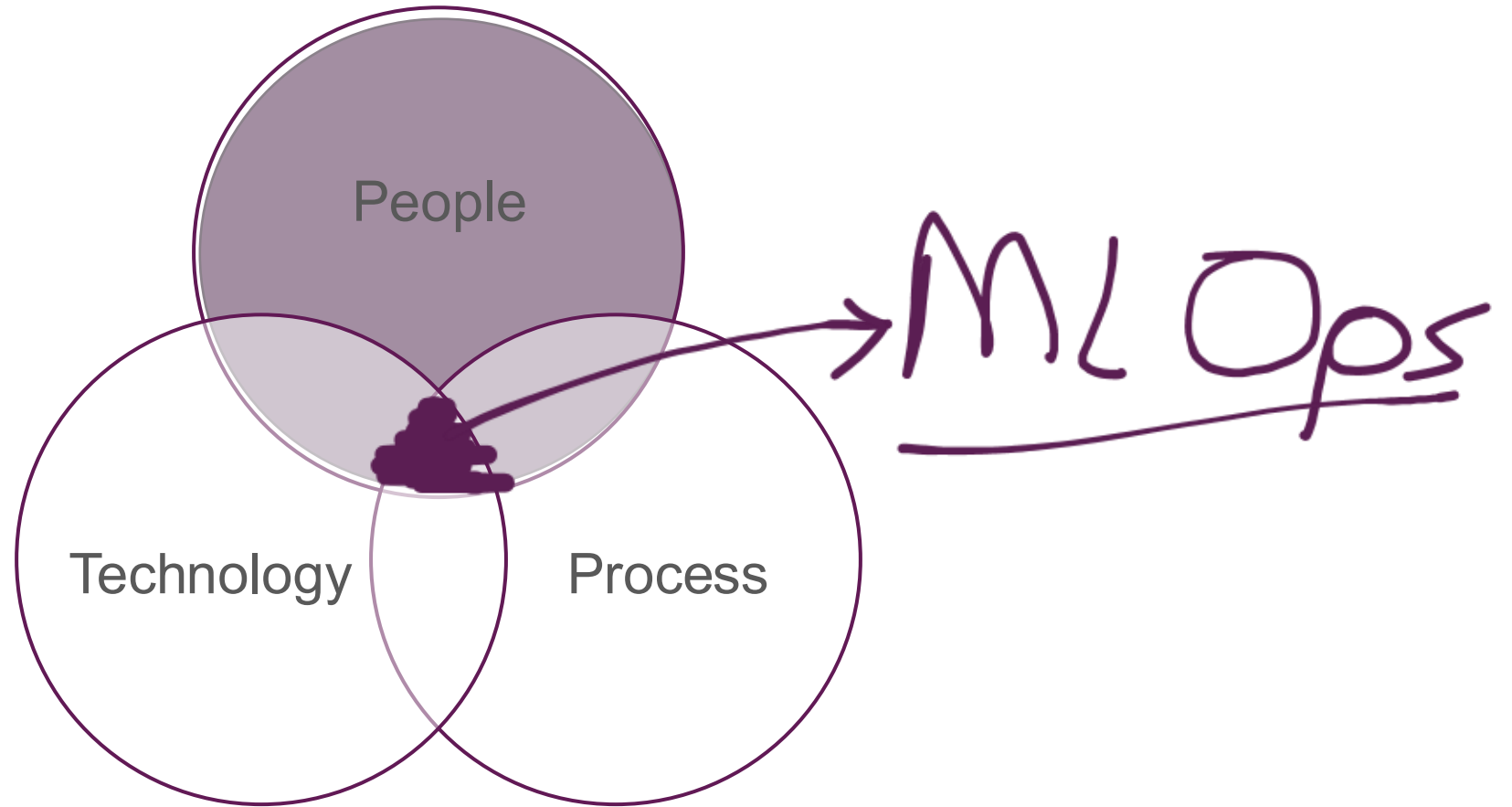
DevOps

Alright so DevOps
is needed as fundament.

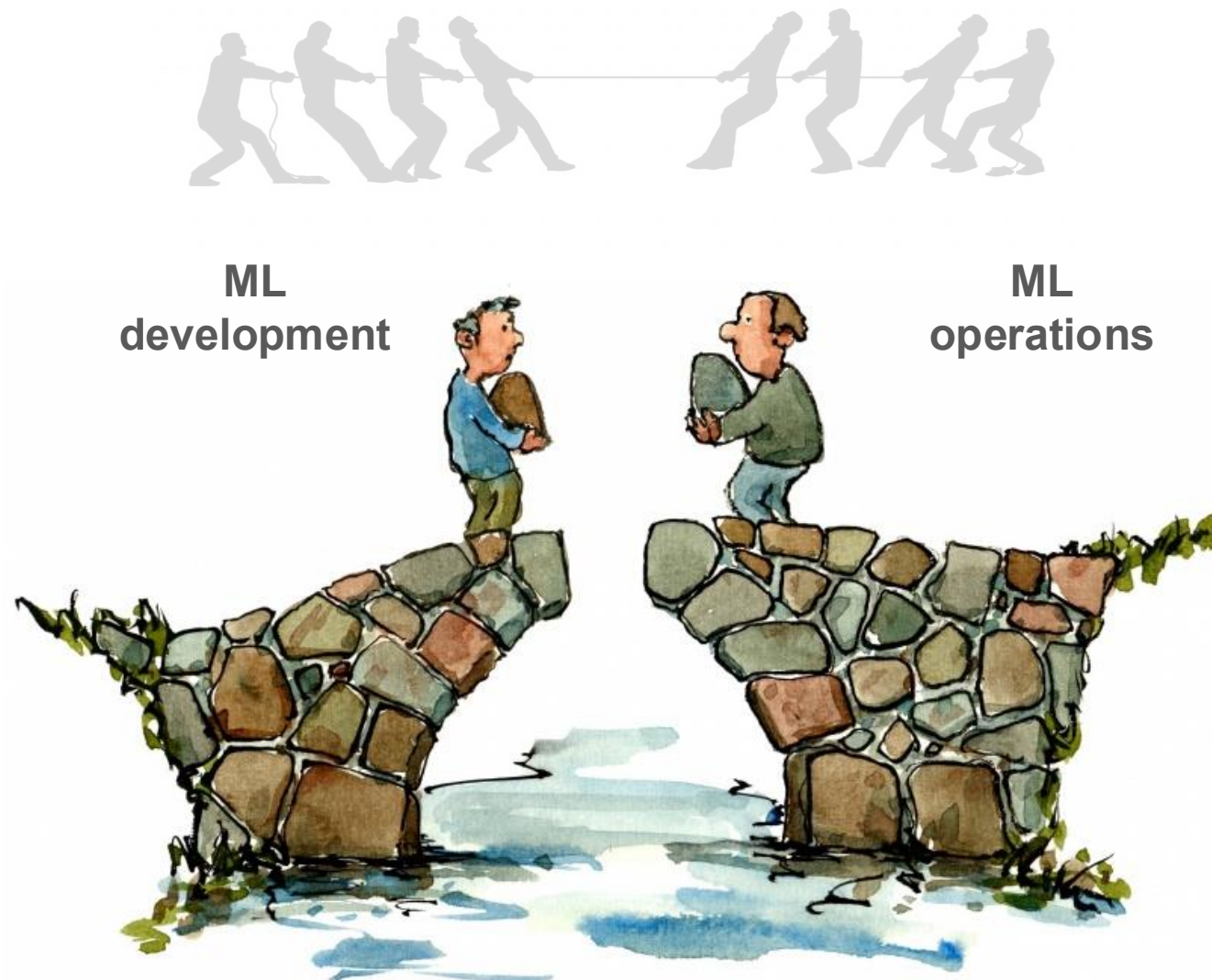
But then what's **the**
MLOps twist?

The MLOps twist

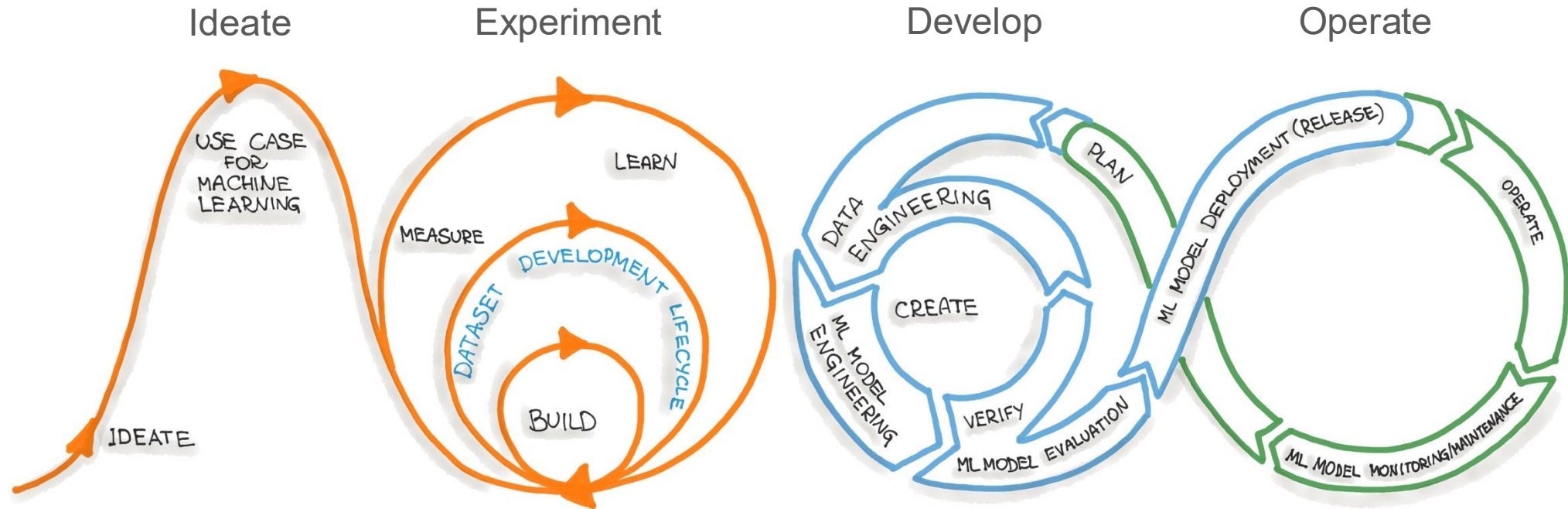
It was never just tech.



The **ML**Ops twist



The end-to-end ML life cycle

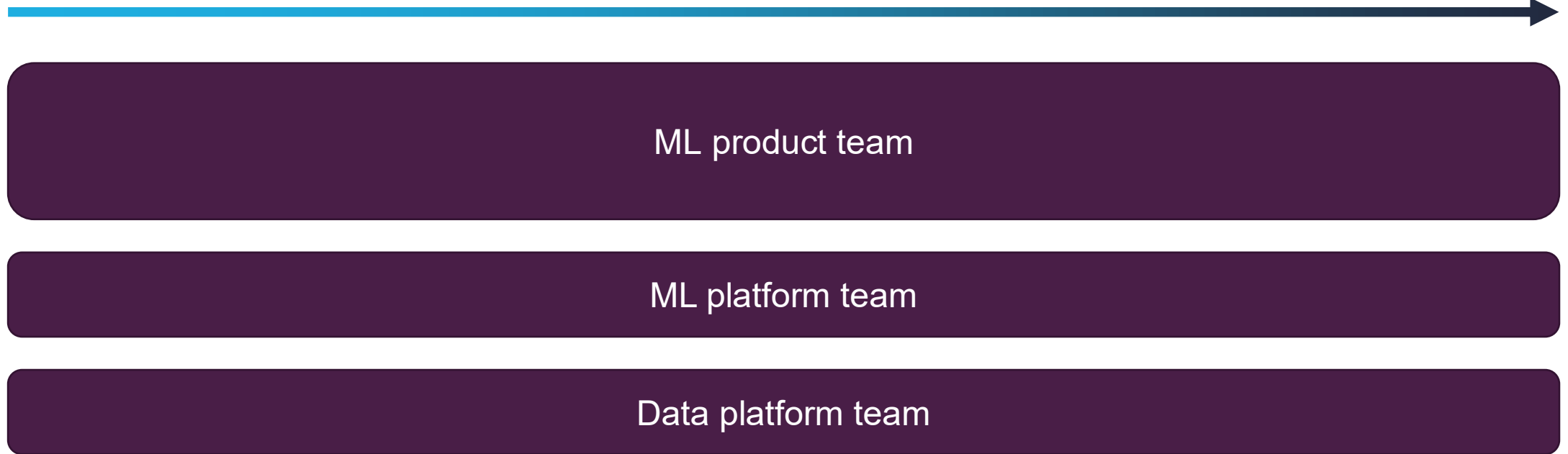


ML Product Team

...and enable end-to-end responsibility

Development

Operations



Measuring the success of MLOps

ML Team – DORA metrics

Deployment frequency

- How does a team release changes to an ML model in production?

Lead time for changes

- How long does it take for changes to get deployed to production?

Change failure rate

- How many deployments cause a failure in production and require intervention?

Time to restoration

- How long it takes to recover from a failure in a production model?

MLOps Team – Platform metrics

Platform adoption

- What percentage of ML teams have become platform users?

User happiness

- Are ML teams happy with the platform services offered?

Capabilities coverage

- What capabilities are supported by the platform?

Cost per user

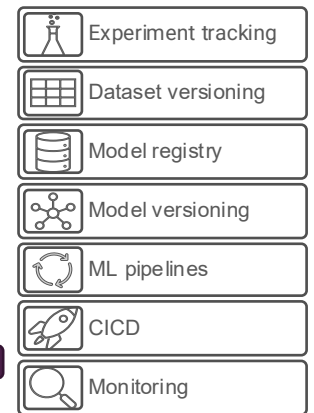
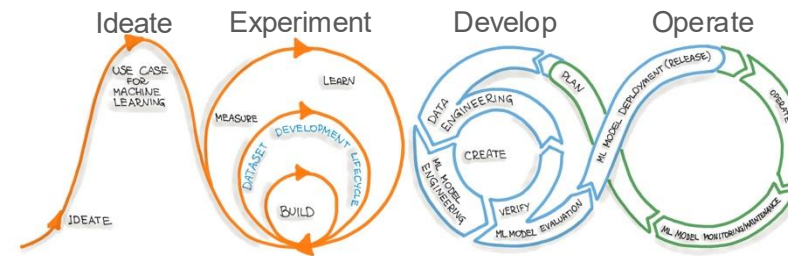
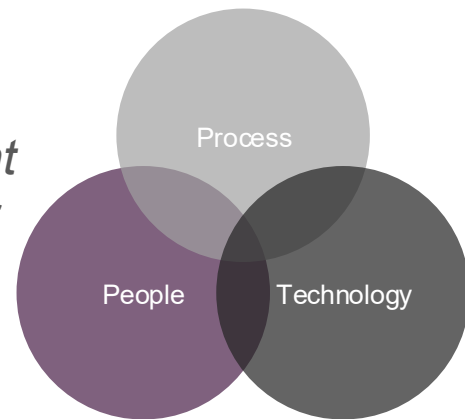
- What is the cost per user, including development and licensing costs?

Build MLOps on top of a strong DevOps foundation & Bring ML Dev + Ops together with end-to-end product teams

MLOps

Bring ML Development & Operations together

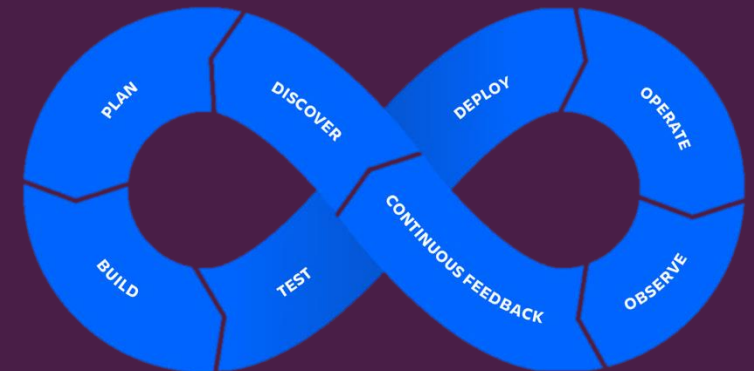
Across all pillars



DevOps

Build on top of the Key DevOps principles

1. Collaborate 🧑‍🤝‍🧑
2. Automation ⚙️
3. Continuous Improvements 📈
4. Customer-centric action 👑
5. Create with the end in mind 🚩





Thank you!
for your attention



Jetze Schuurmans
Machine Learning Engineer
jetze.schuurmans@xebia.com



Jeroen Overschie
Machine Learning Engineer
jeroen.overschie@xebia.com

