Are you ready for **MLOps**? 👉

Jetze Schuurmans
Jeroen Overschie

PyData London — 16th of June 2024

Xebia | Data

Jeroen Overschie

Machine Learning Engineer

Jetze Schuurmans

Machine Learning Engineer

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

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

Why MLOps?

Why MLOps?

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

Long development cycles

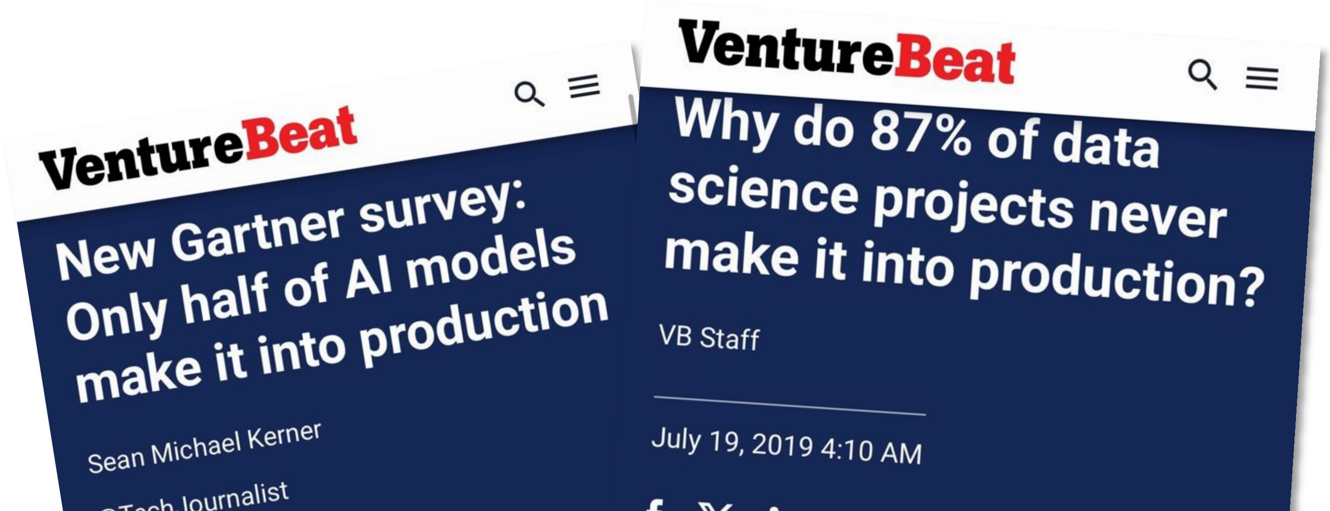
To create models, update models or add features

Lacking feedback

On model performance and added value

Few models in production

Unreliable solutions



Why MLOps?

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

🤷‍♀️🤷‍♂️🤔

How come?

Why is this still difficult these days?

Long development cycles

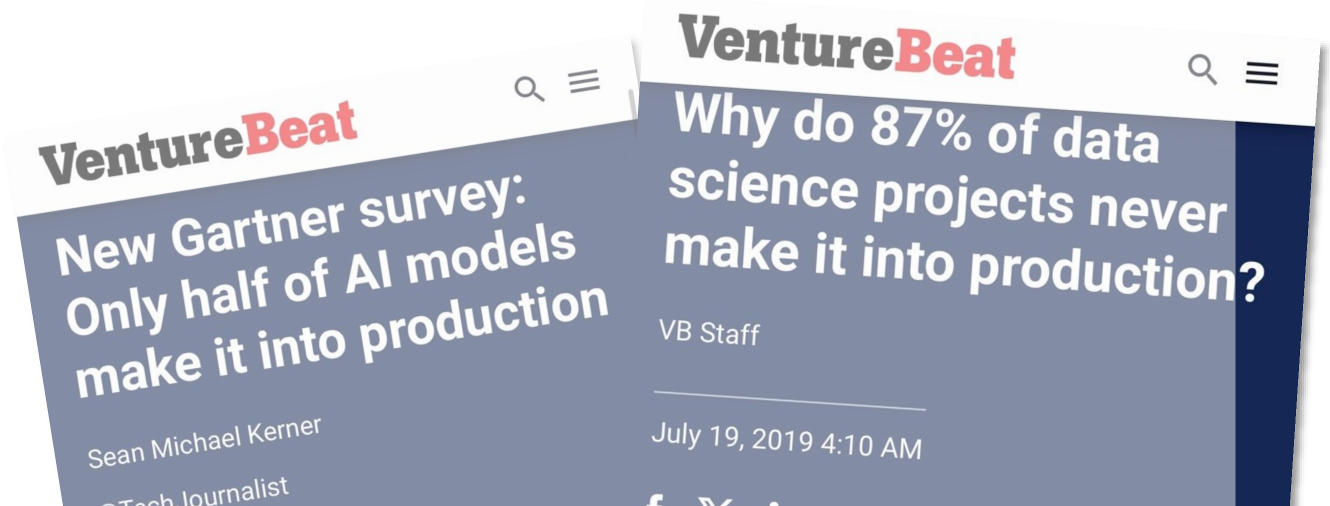
To create models, update models or add features

Lacking feedback

On model performance and added value

Few models in production

Unreliable solutions



VentureBeat

New Gartner survey: Only half of AI models make it into production

Sean Michael Kerner

@Tech Journalist

VentureBeat

Why do 87% of data science projects never make it into production?

VB Staff

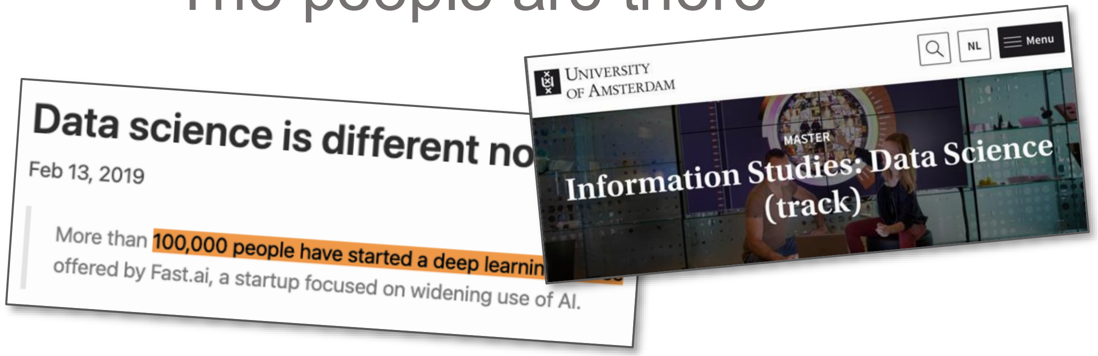
July 19, 2019 4:10 AM

The tech is there

How come?

Why is this still difficult these days?

The people are there

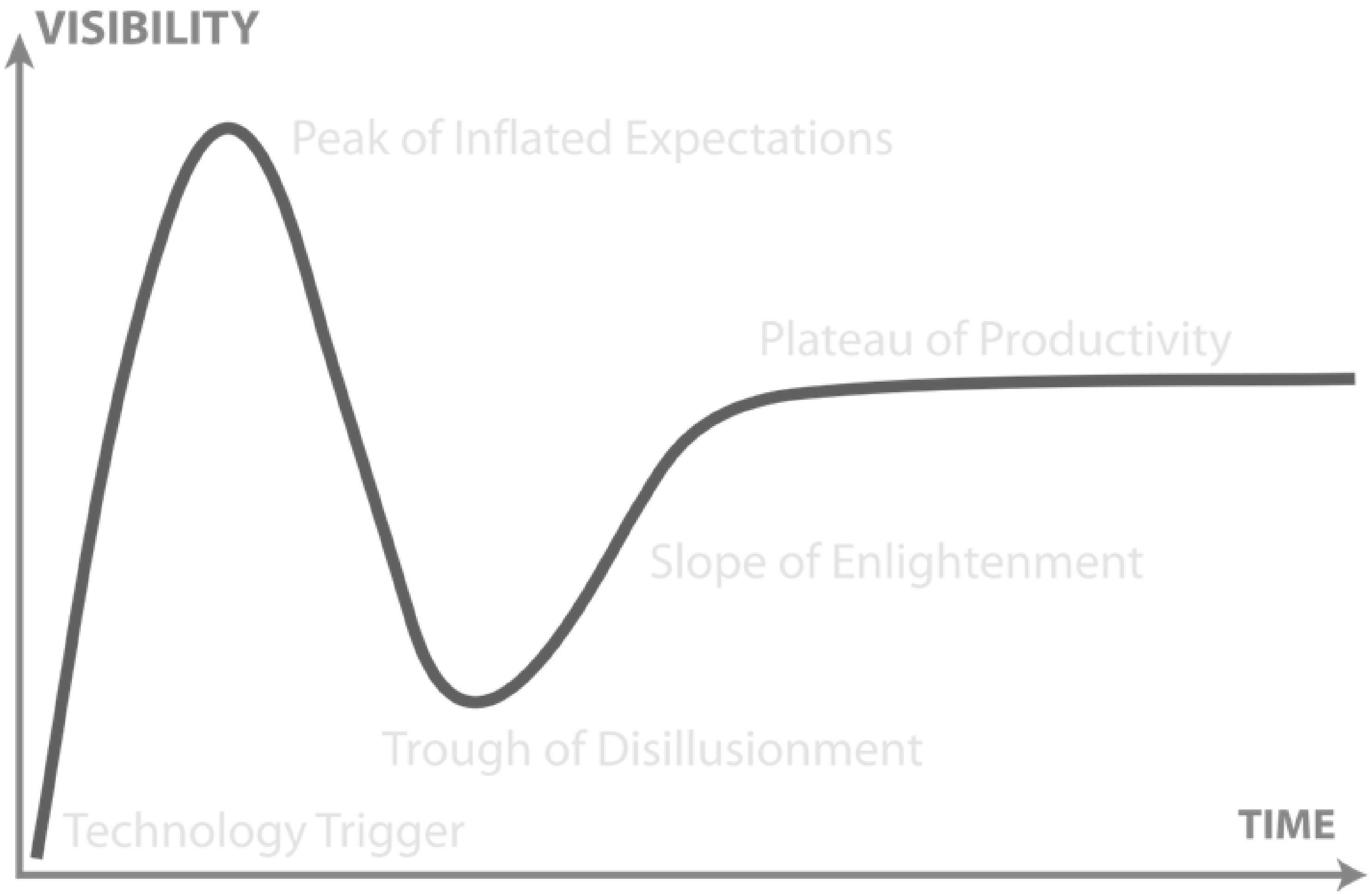


The meetups are there 😉

The promise of MLOps

Hype cycle

The promise of MLOps

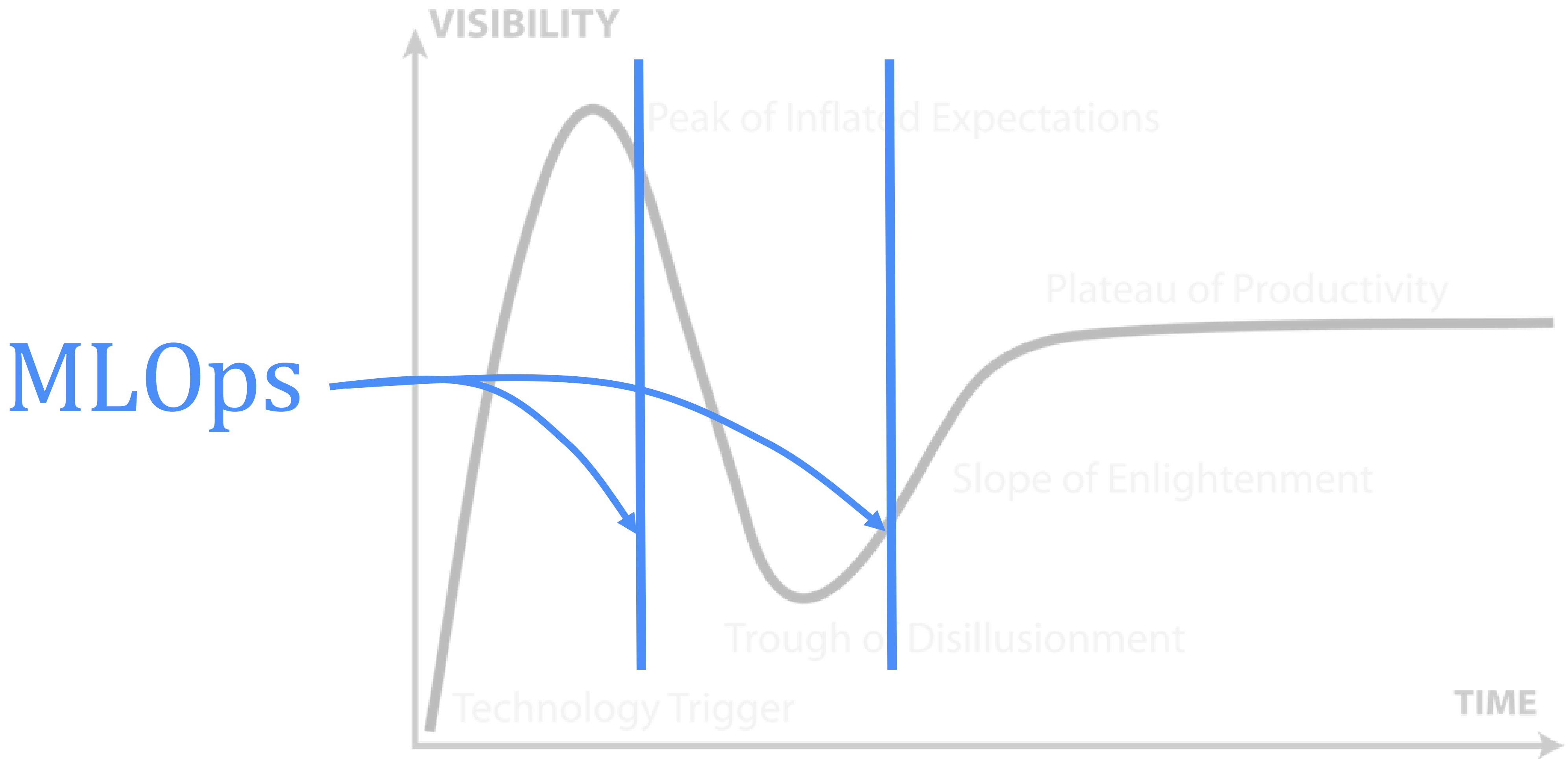


[1] https://en.wikipedia.org/wiki/Gartner_hype_cycle

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]

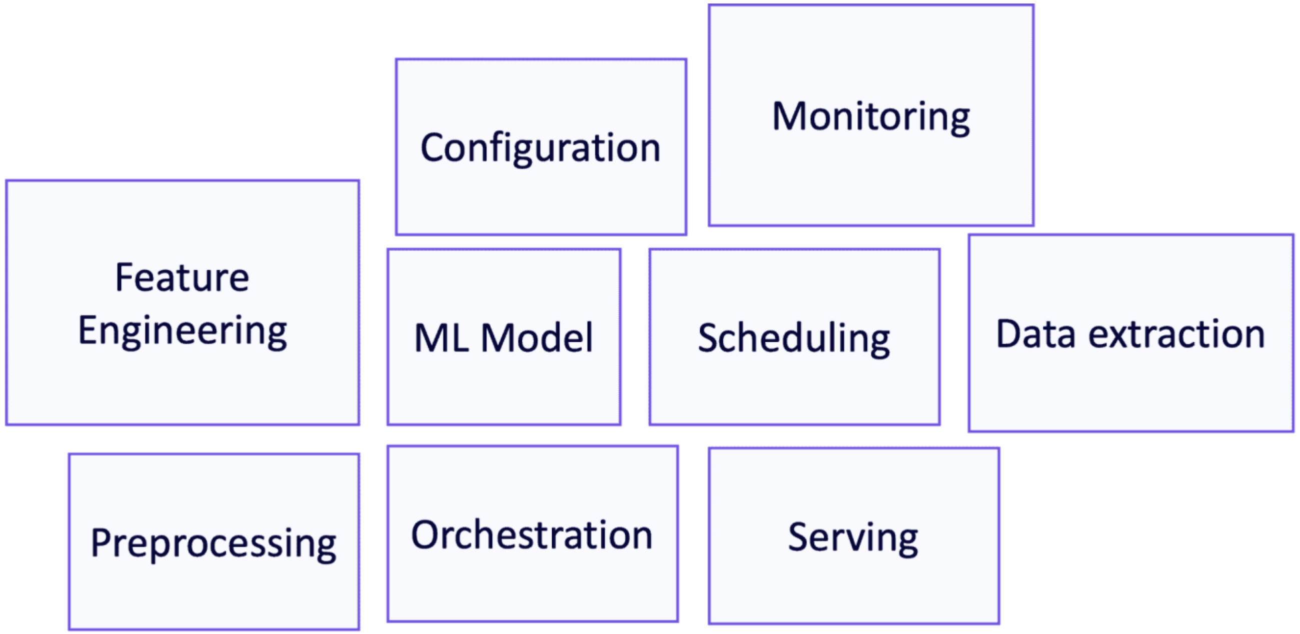
The promise of MLOps

Hype cycle



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

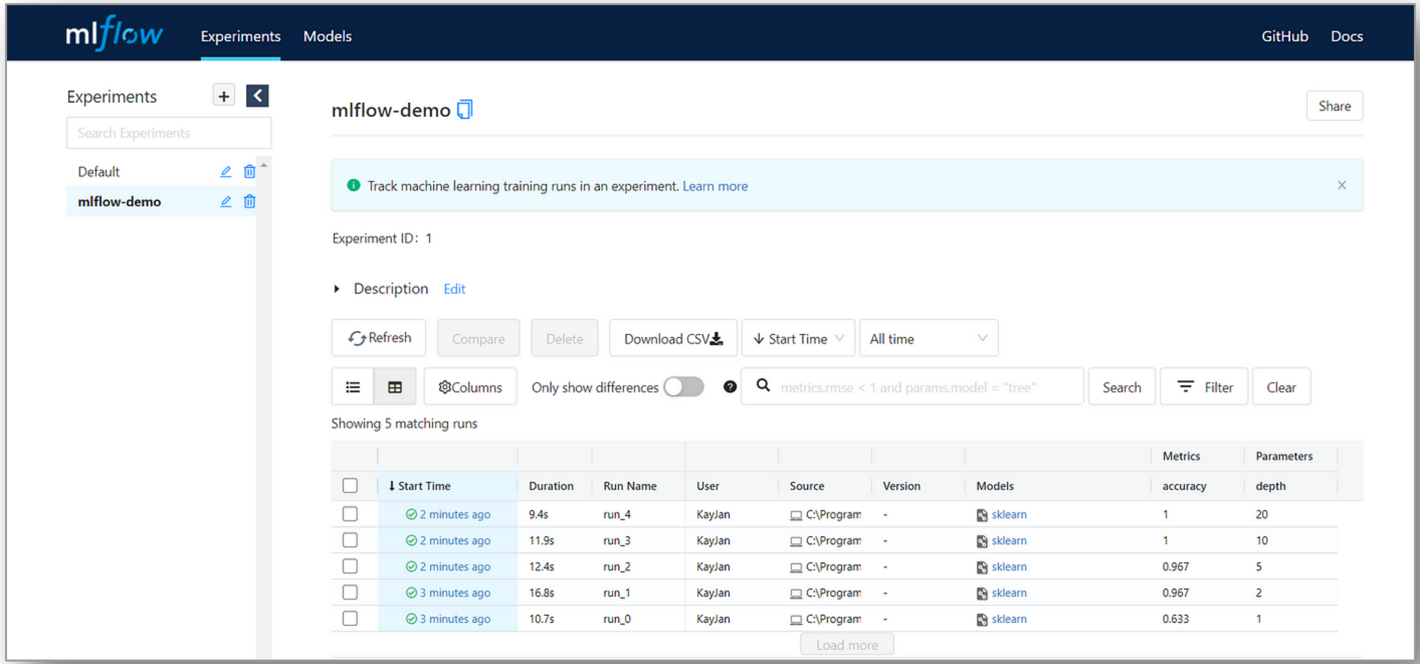
The promise of MLOps



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

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

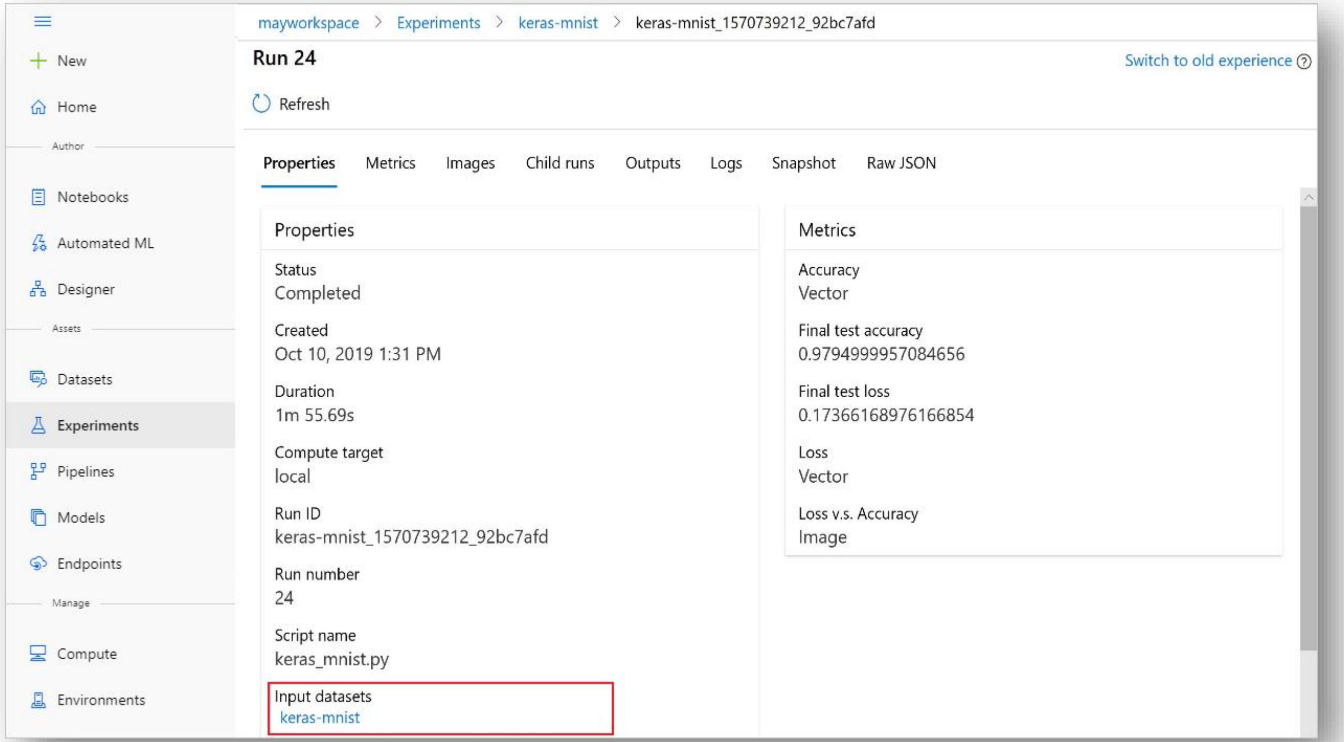
The promise of MLOps



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

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

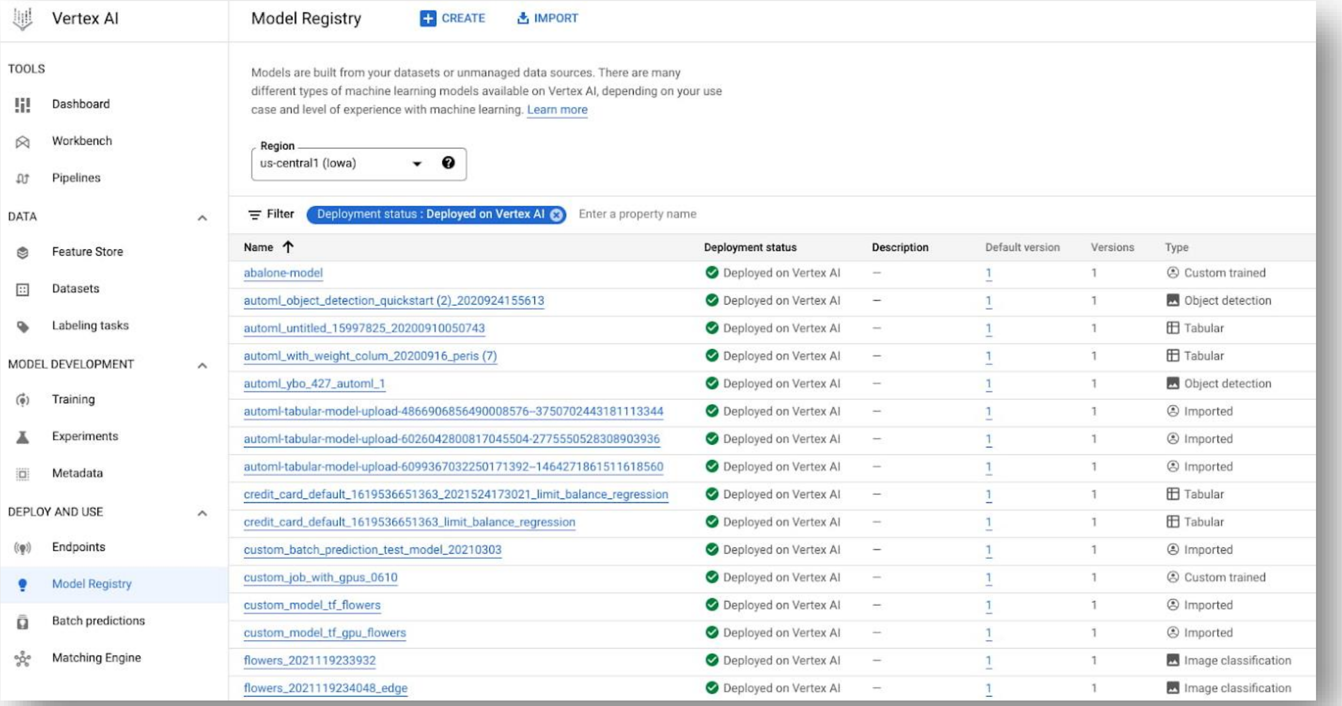
The promise of MLOps



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

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

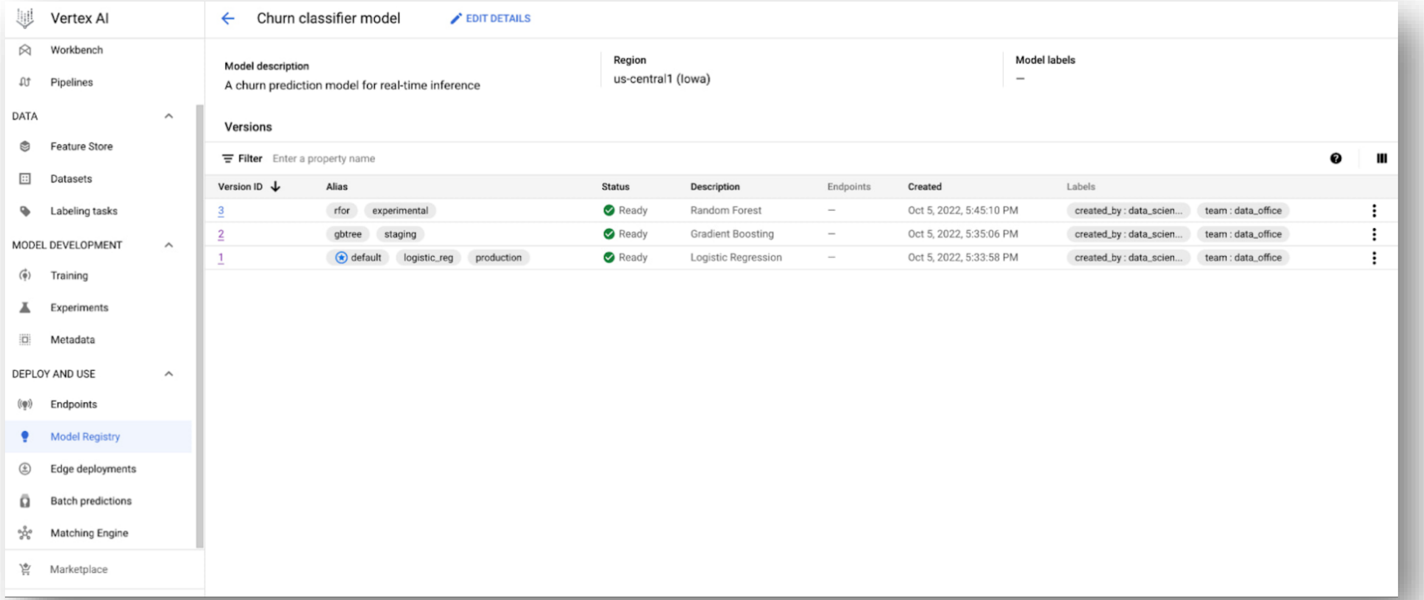
The promise of MLOps



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

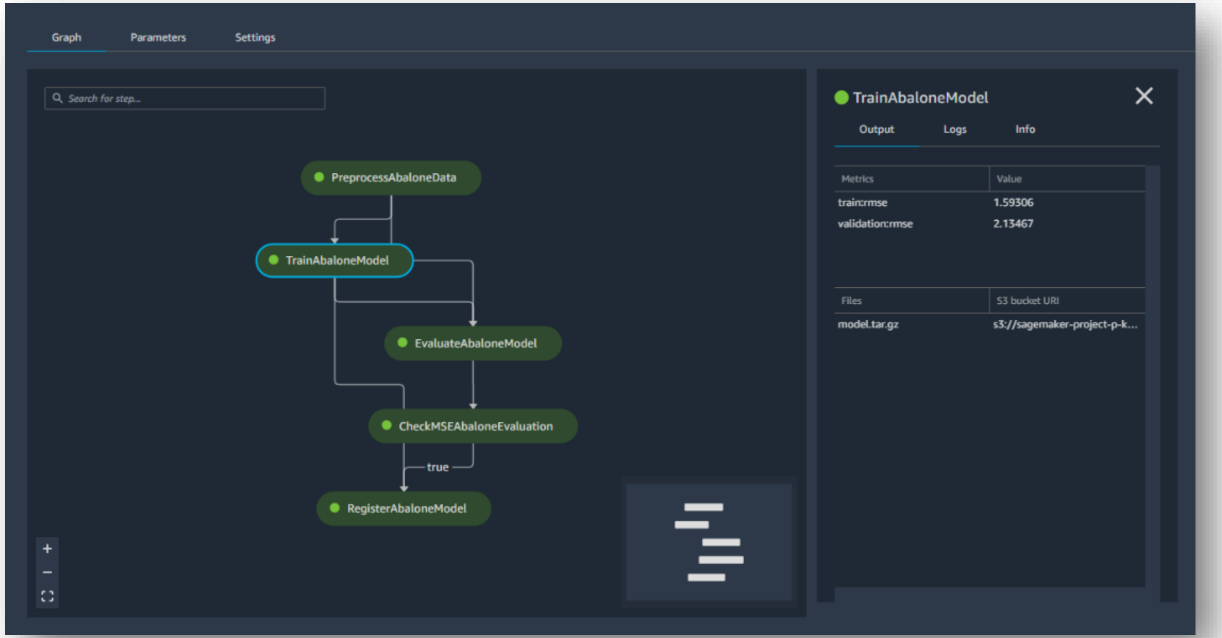
The promise of MLOps



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

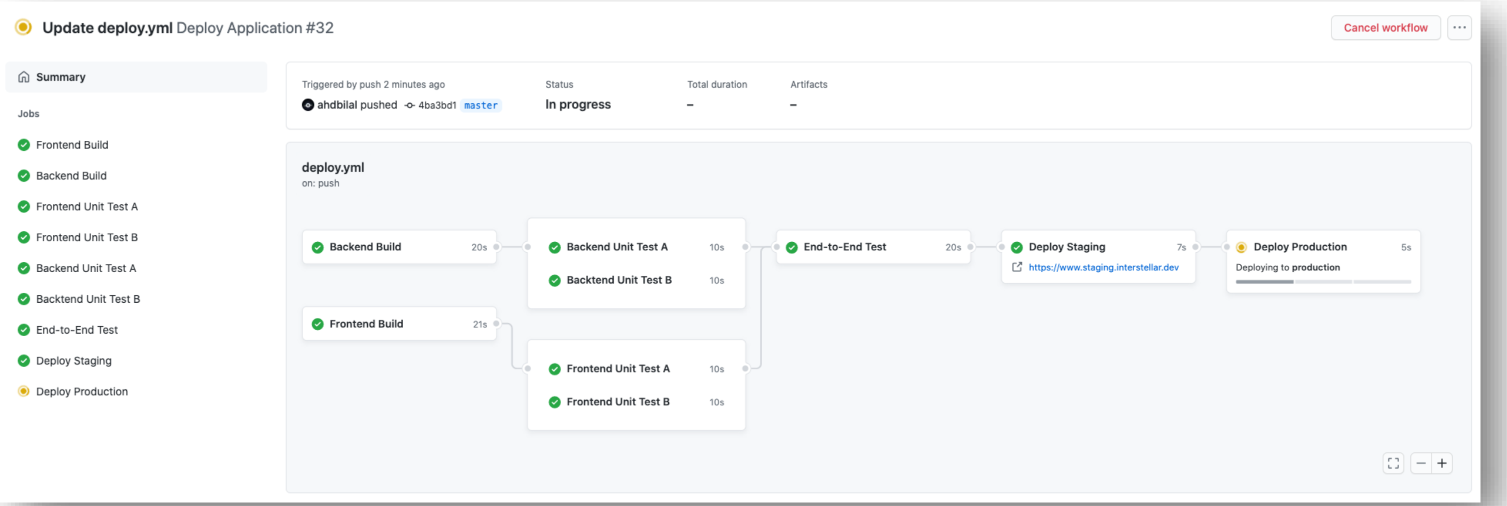
The promise of MLOps



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

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

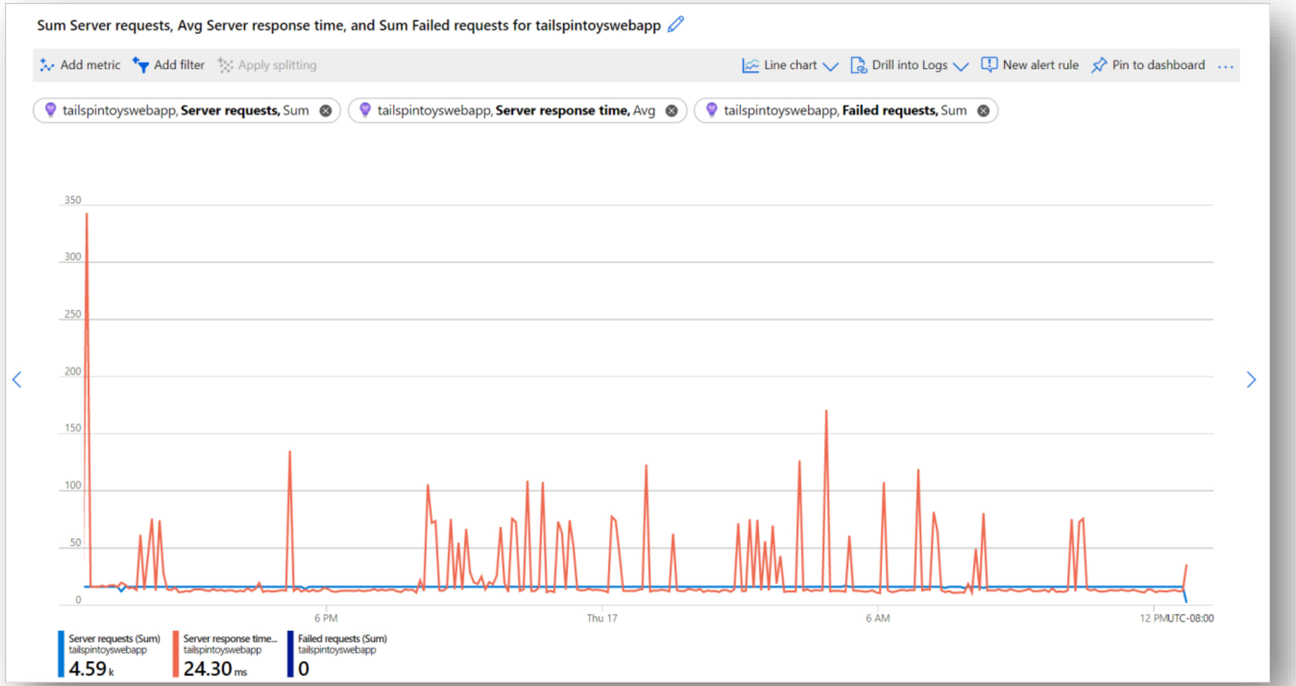
The promise of MLOps



https://github.blog/changelog/2020-12-08-github-actions-workflow-visualization/

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- **CICD**
- Monitoring

The promise of MLOps



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

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

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**.

That sounds great but,

Where's this all come from?

Long development cycles

To create models, update

Lacking feedback

On model performance

Few models in production

Unreliable

MLOps promises

Speed-up development with automation and collaboration

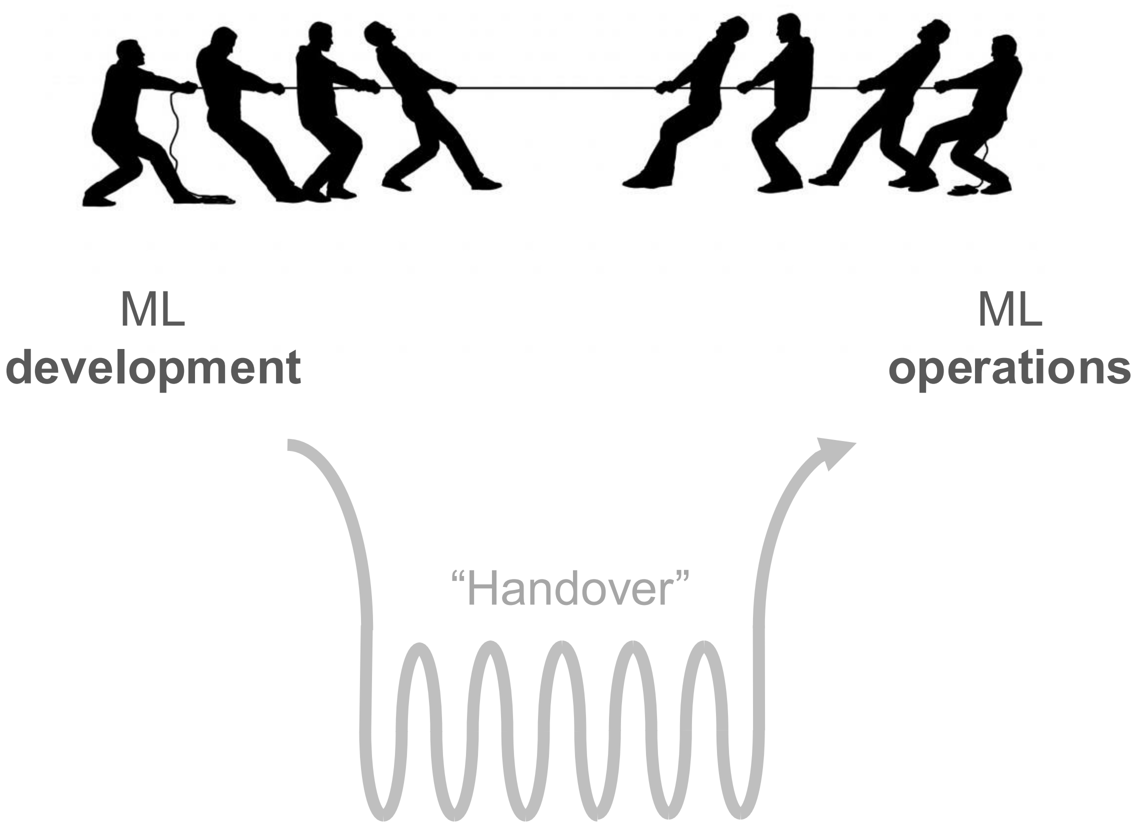
Focus on rapid feedback and iterative development

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**

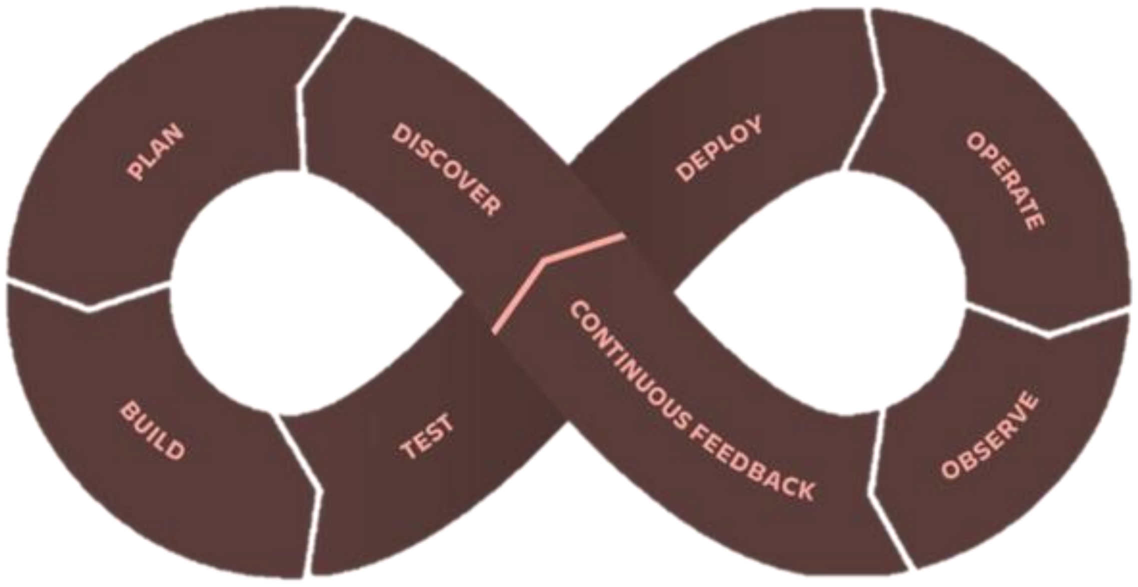


A thing called

DevOps

Since 2007

Bring together Dev and Ops

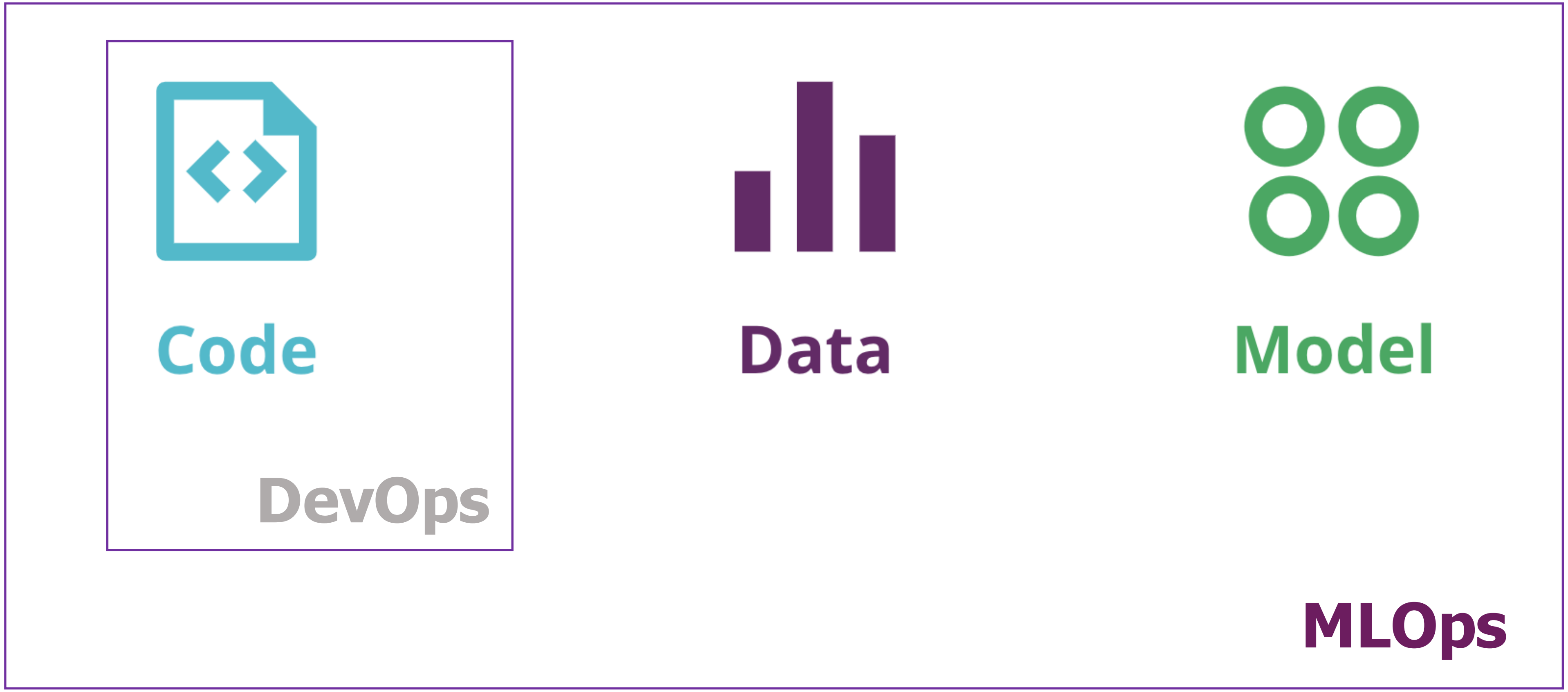


5 key principles:

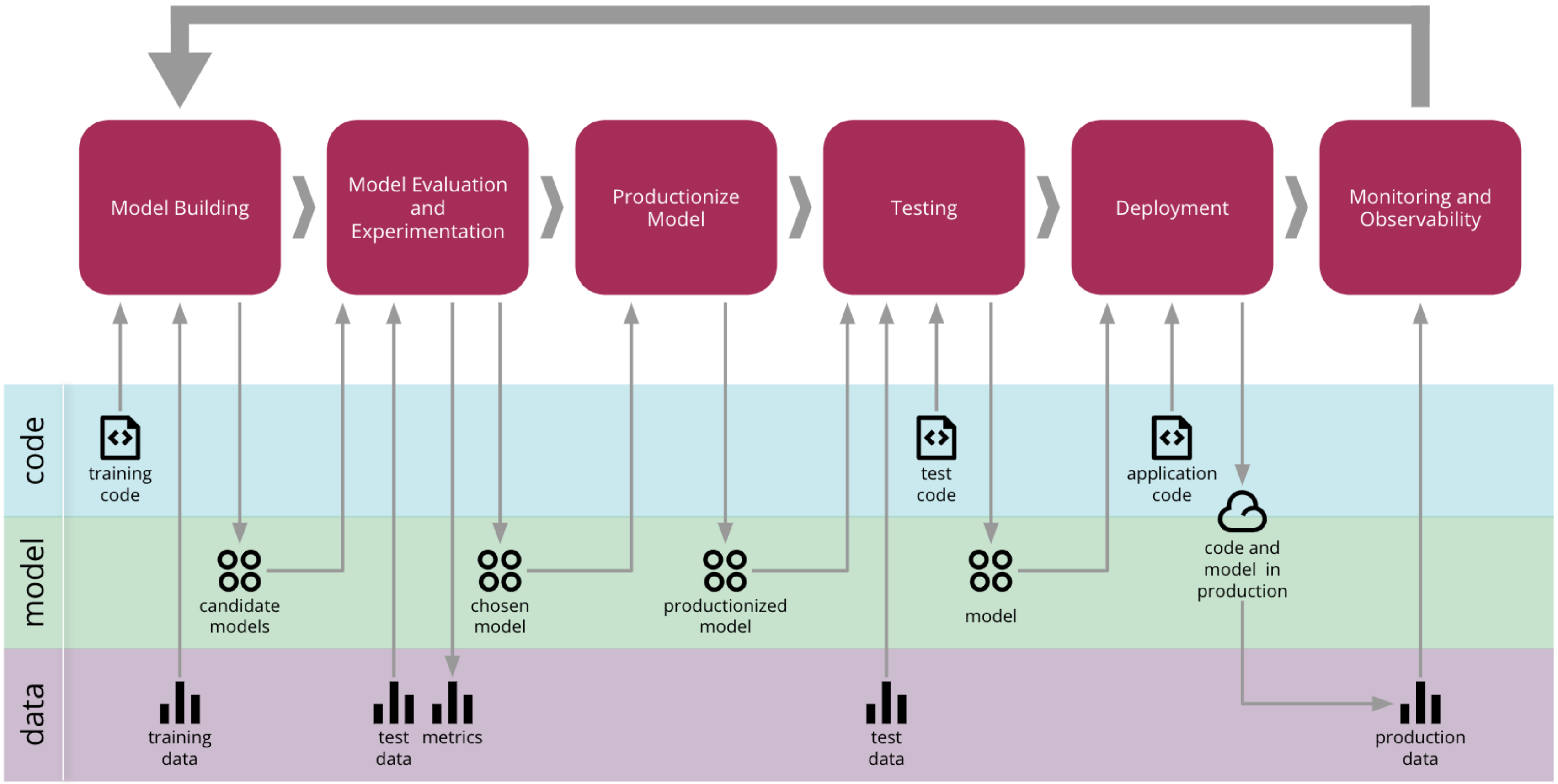
1. Collaborate 🤝
2. Automate ⚙️
3. Continuous Improvements 📈
4. Customer-centric action 👑
5. Create with the end in mind 🏁

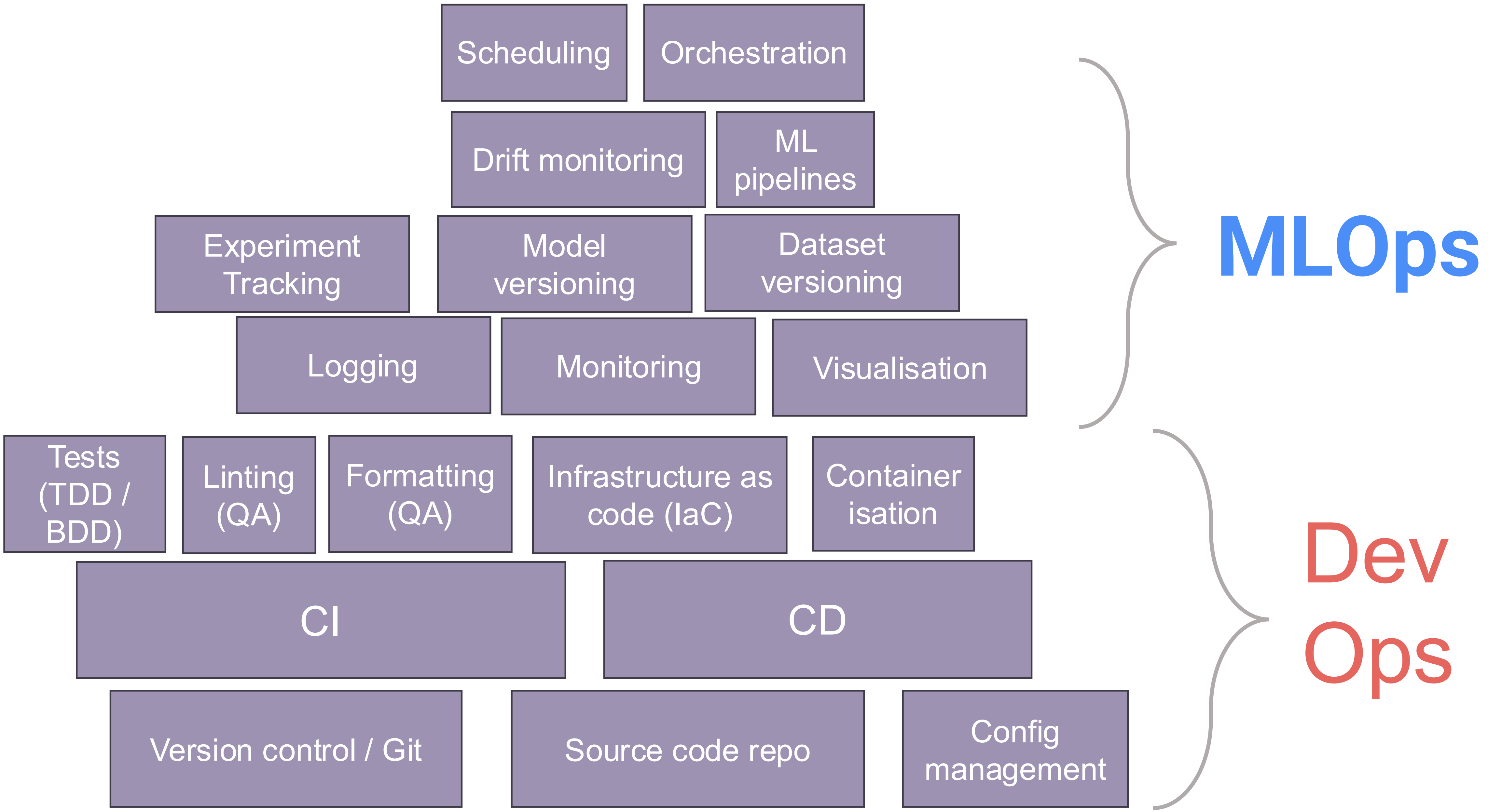
Source: https://www.atlassian.com/devops/what-is-devops

DevOps vs. MLOps



DevOps & MLOps
You need both in a Machine Learning Pipeline




Scheduling
Orchestration
Drift monitoring
ML pipelines
Experiment Tracking
Model versioning
Dataset versioning
Logging
Monitoring
Visualisation
Tests (TDD / BDD)
Linting (QA)
Formatting (QA)
Infrastructure as code (IaC)
Container isation
CI
CD
Version control / Git
Source code repo
Config management
MLOps
Dev Ops

MLOps

Dev
Ops

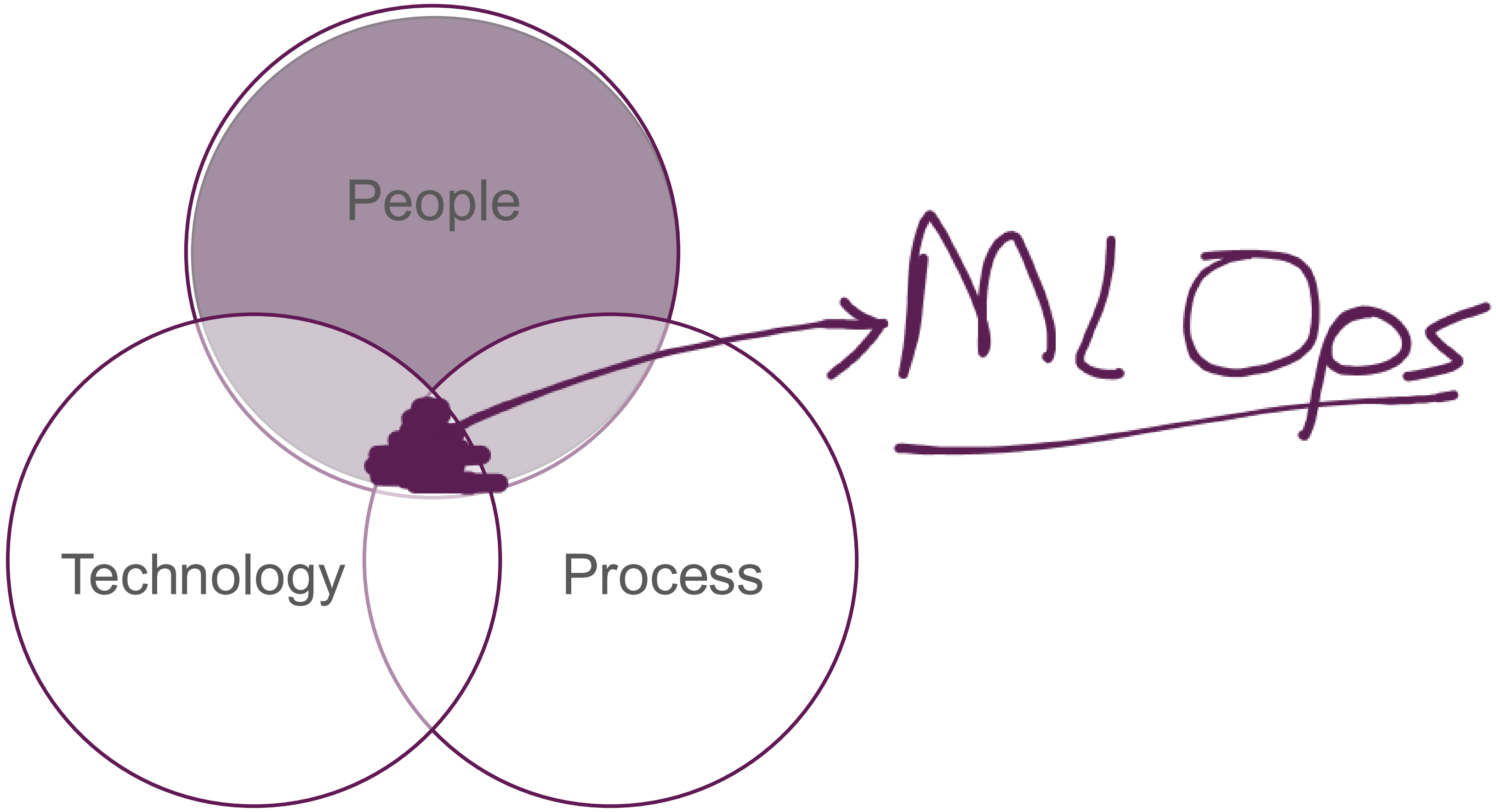
DevOps

Alright so DevOps
is needed as fundament.

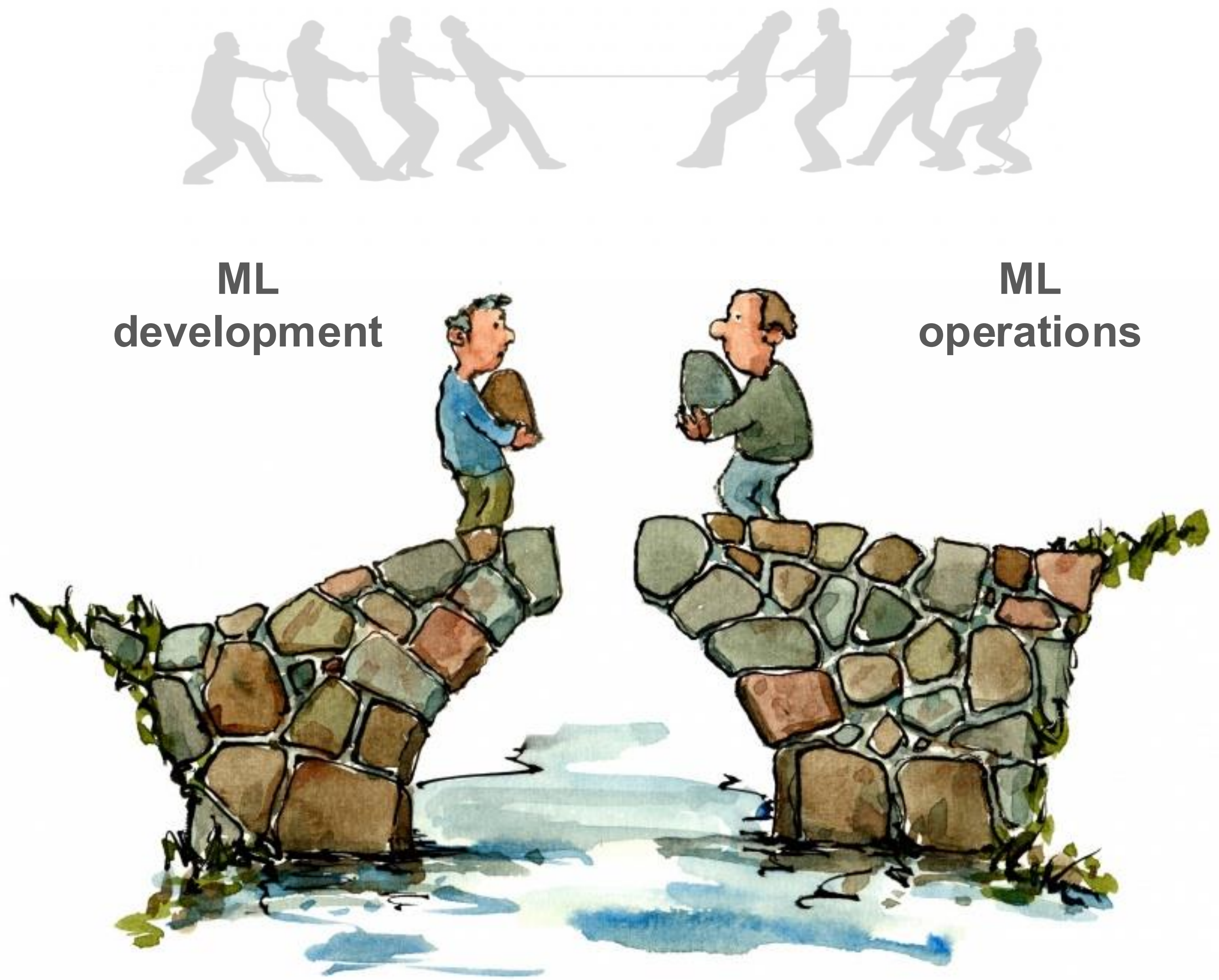
But then what’s **the**
MLOps twist?

The MLOps twist

It was never just tech.



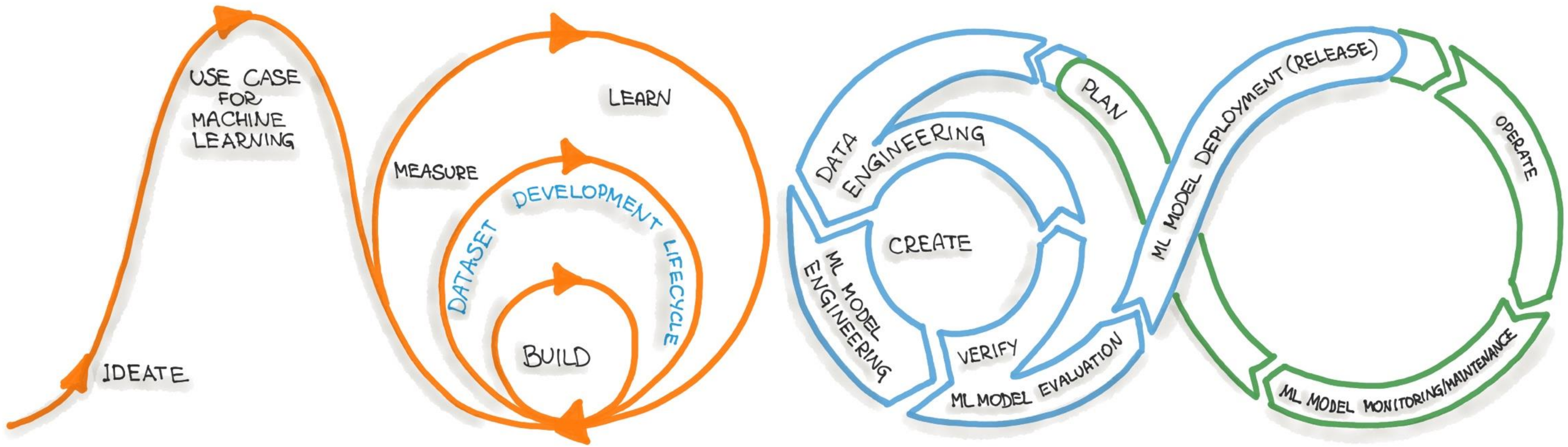
The MLOps twist



Adapted from www.thequotablecoach.com

The end-to-end ML life cycle

Ideate | Experiment | Develop | Operate



ML Product Team

...and enable end-to-end responsibility

Development → Operations

ML product team

ML platform team

Data platform team

Xebia

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 cahnges 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?

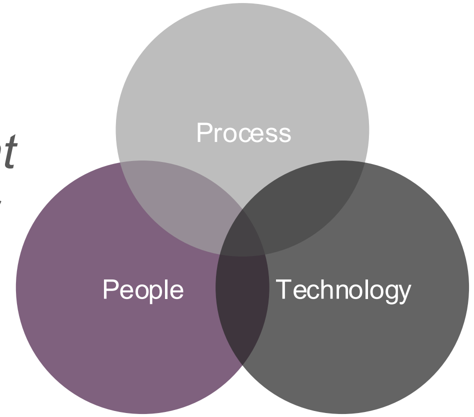
Source: xebia.com/blog/how-to-measure-your-mlops-performance/

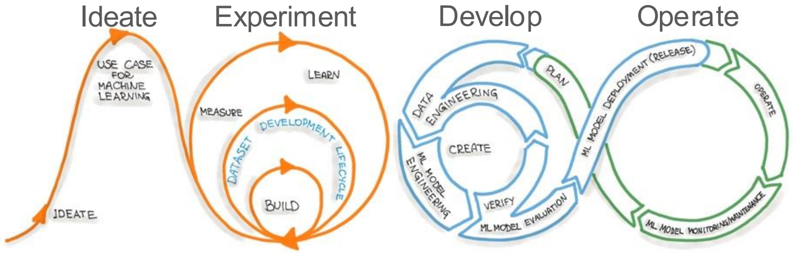
Build MLOps on top of a strong DevOps foundation & Bring ML Dev + Ops together with end-to-end product teams

MLOps

Bring ML Development & Operations together

Across all pillars





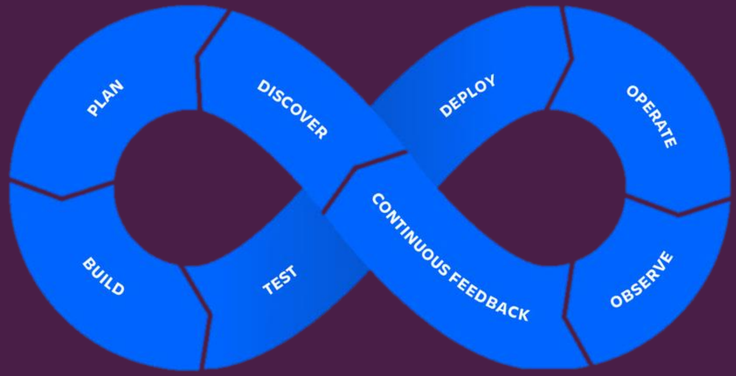
ML Product Team

- Experiment tracking
- Dataset versioning
- Model registry
- Model versioning
- ML pipelines
- CICD
- Monitoring

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 🏁



Xebia | Data

PyData GB

Thank you!
for your attention

Jetze Schuurmans

Machine Learning Engineer

jetze.schuurmans@xebia.com

Jeroen Overschie

Machine Learning Engineer

jeroen.overschie@xebia.com