• Live webinar

Empowering Data Scientists with MLOps

Tuesday, Jan 26th @ 12pm EST (9am PST)

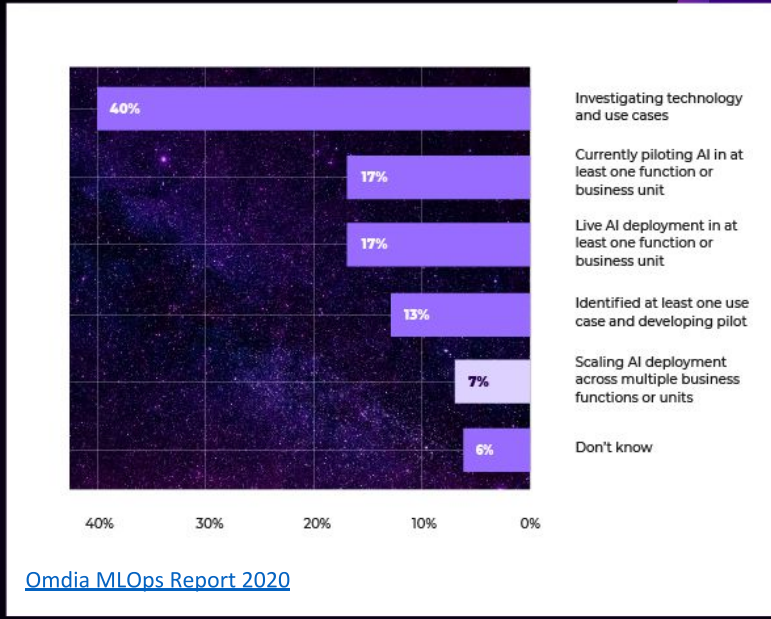
cnvrg.io

// whoami

- Developer/Data Scientist → CEO
- cnvrg.io - built by data scientists, for data scientists to help teams:
  - Manage, build and automate ML from research to production
  - Bridge science, engineering, and infrastructure

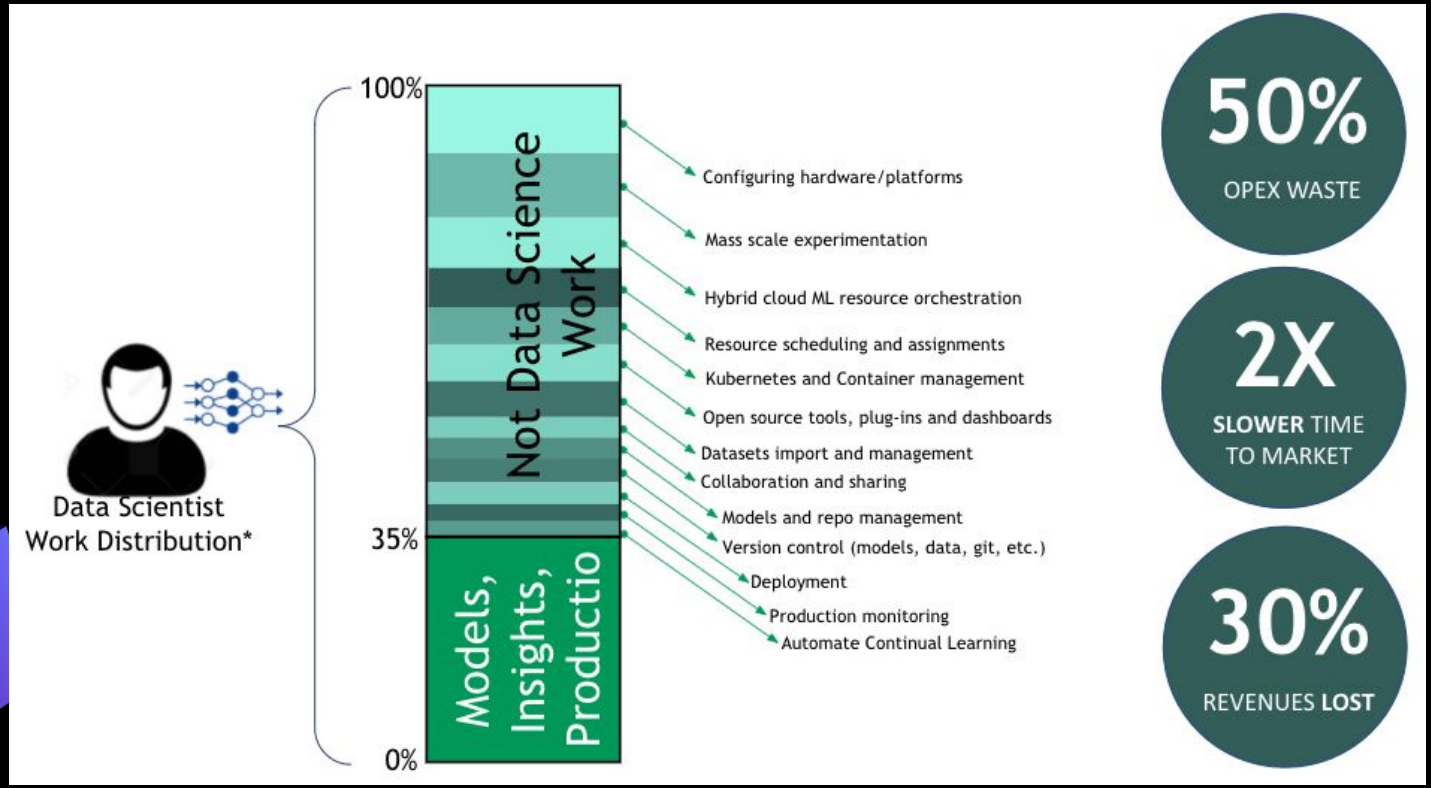
Industry Problem

- 80% of enterprises fail to scale AI
- Data scientists spend >65% of time on non-data science tasks
- Huge gap between isolated experimentation and widespread implementation
- Enterprise AI practitioners suffer from the same challenges software developers faced back in the 1990

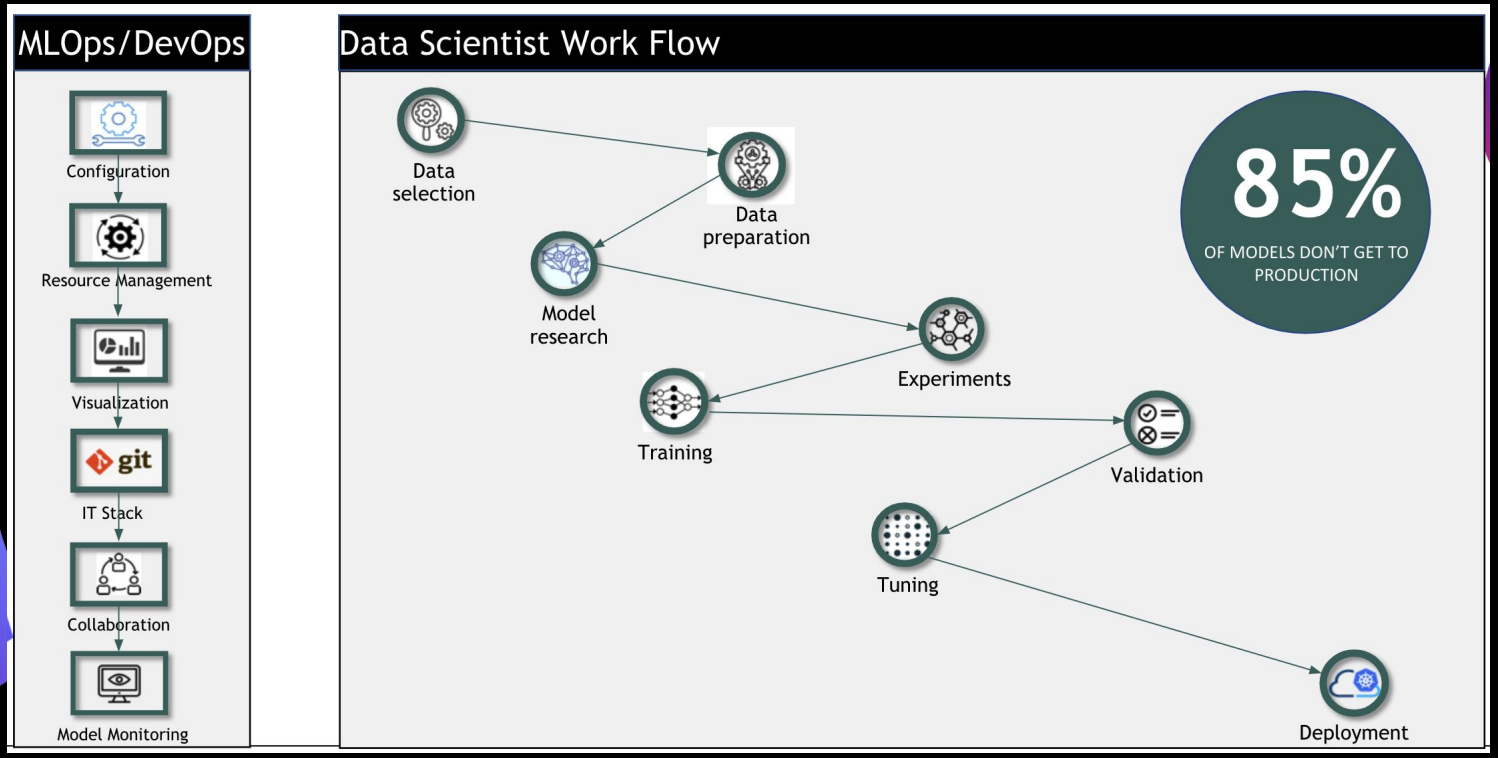


Omdia MLOps Report 2020

// Data Scientists doing little DS work

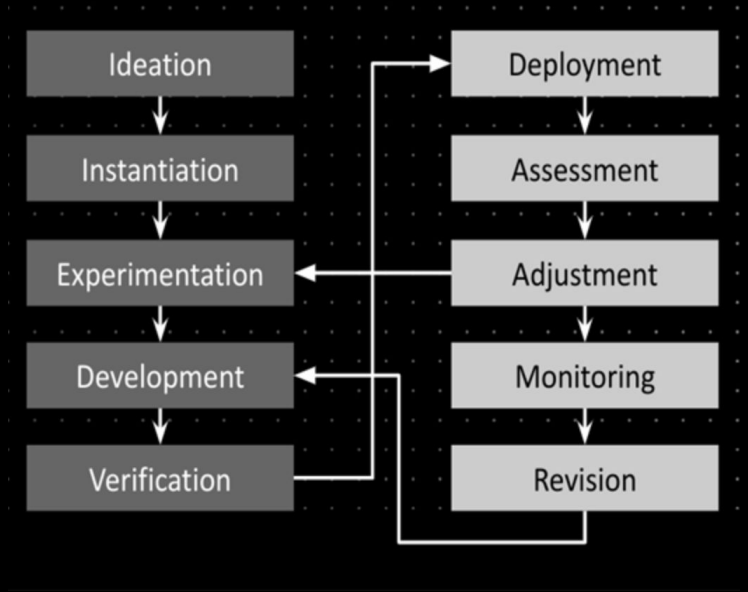


// ML Workflows are siloed

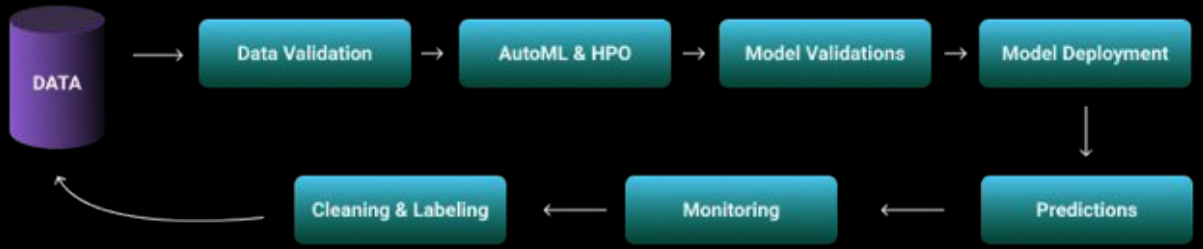


What is MLOps?

- Enables data engineers, data scientists, and IT professionals to work collaboratively while still focusing on their respective ML project tasks
- Bridge the gap between the experimental aspects of data science and the operational nature of software deployment, creating a living lifecycle
- Key MLOps automations: data versioning, code versioning, deployment orchestration, centralized artifact repositories, CI/CD & monitoring in production

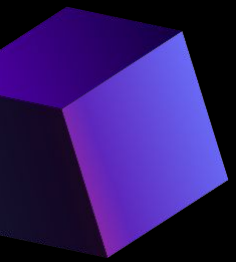


An ML Pipeline



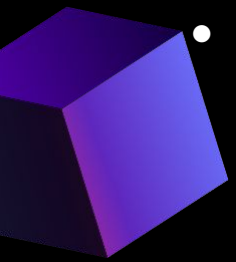
Building your own MLOps solution

- Containers, Kubernetes is the standard
- Data Scientists != DevOps/ML Engineers
- Different hardware requires different configuration: CPU/GPU
- MLOps for Computer Vision is similar to MLOps for analytics, but can also be very different
- Integrate to the existing IT stack



MLOps for Research

- Rapid experimentation, optimizing models, running research, implementing papers without DevOps slowing you down
- Launch workspaces and IDEs on any compute resource instantly: GPU, CPU clusters within reach
- Track and monitor your workloads with model tracking and experiment tracking
- Collaboration on datasets, optimizing data access
- Build reports, dashboards, web apps for better knowledge sharing and collaboration



MLOps for Production

- ML Pipelines: orchestrate training based on triggers:
  - new data? Schedule? Model performance?
- Serving models: easily deploy models as APIs/stream/batch with support of different frameworks and the ability to update model versions, A/B testing
- Monitor pipelines and models in production: track behavior of pipelines (data quality, model quality) and models in production (prediction performance, bias, feature validation)
- Fast debugging and transition to research