WORKSHOP DAY ONSITE
OCTOBER 4, 2023

CONFERENCE DAY ONSITE
OCTOBER 5, 2023

WORKSHOP DAY ONLINE
OCTOBER 6, 2023

ML at Scale.

How to productionize and operationalize hundreds of ML solutions in a simple and reliable way?

Tom Deregowski
Director of AI/ML Engineering, Nike

SUMMIT.DATAMASS.IO

Challenges with scaling up ML on the enterprise level

10%

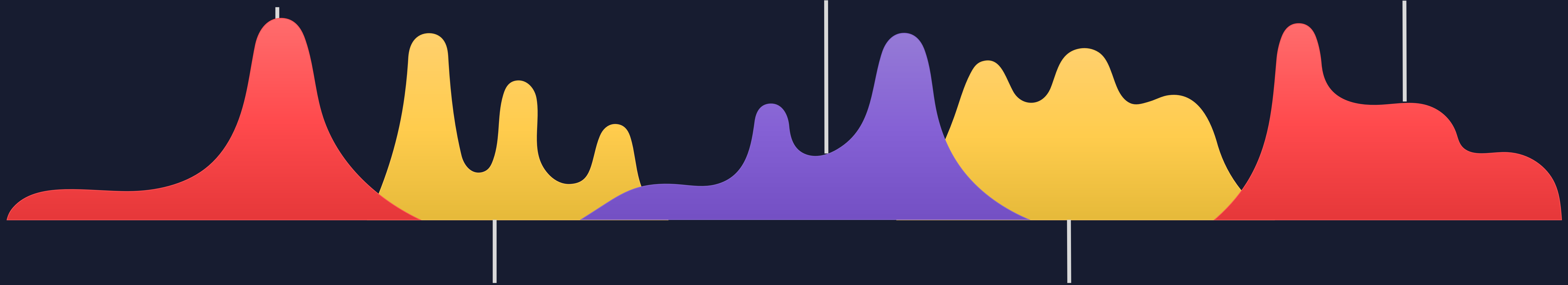
Models that made it to production in the organizations where data culture is in development [1]

97%

It takes **more than** one **month** to **obtain** the necessary **datasets** needed for analysis [3]

92%

It takes more than three months for a model to get into production [2]



75%

Models that get into **production**, need **more than one week** to do it [3]

63%

Models **need IT support** for exposing REST APIs, orchestrion, refactoring and integration with AIML products [3]

[1] Top Strategic Technology Trends, Gartner 2020
[2] Accelerating AI deployments — paths of least resistance, Gartner 2022
[3] Internal research, 2021

ML Engr. is more complex than Software Engr.

Artefacts

- Code
- Data (Data Drift)
- Model (Model Drift)

Processes

- Collect & integrate the data
- EDA & Feature Engineering
- Develop & Tune
- Train & Evaluate
- CI/CD
- Model Management
- Monitor App, the data & model
- Online Experimentation
- Feedback loop
- Continuous training

Roles

- Business Stakeholder
- Platform Engineer
- Software Engineer
- DevOps Engineer
- Data Steward
- Data Engineer
- Data Scientist
- ML Engineer

Tools

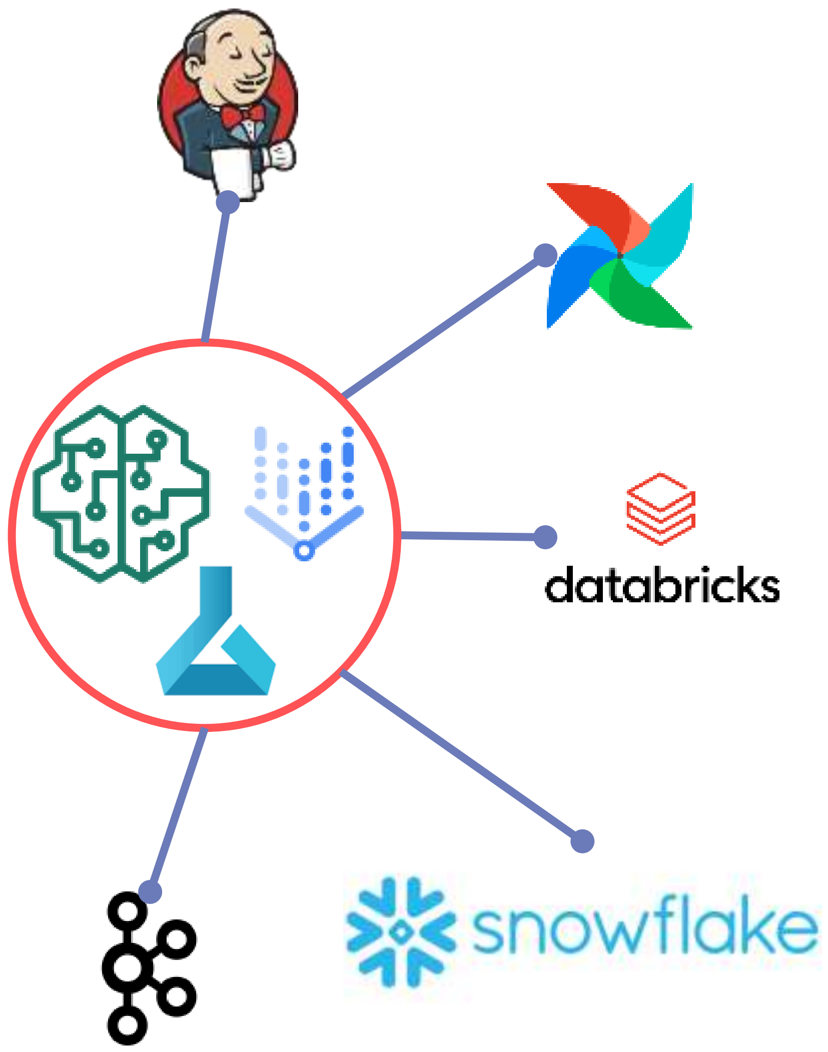
- IDE
- DS Notebooks
- Data Science Platforms
- Enterprise ML Platforms
- Data Generating and Labeling
- CI/CD
- MLOps Solutions
- Runtime: Cloud, AI Hardware / GPU
- Computer Vision
- Speech
- NLP

The job to be done

Get from here...

In-house developed solutions

...to here



Guiding Principles

Provide vanilla experience

Do not change the default ML Enterprise Platform experience. Use the tool as Cloud-provider designed it, and ensure that Platform documentation and examples remain relevant.

Leverage Cloud-provider offering

Do not reinvent the wheel; instead, leverage Cloud-provider offering. Treat building new software as a last resort.

Prepare blueprints & bootcamps

Capture common DS and ML Engineers problems and best practices and solutions to them. Codify and automate the best solutions, provide reference implementation and bootcamps. Make it easy to do the right thing in the right way.

MLOps-as-a-service

Automate processes to improve productivity, lower the risk of human error, and enable a self-service

Build the platform for multiple types of users

Build a platform for a diverse group of users: from non-technical DS without engineering experience to ML engineers who are experts in software craftsmanship.

Some customizations are always needed

User onboarding

Integration with corporate standard for seamlessly access requests, access provisioning, access certification and separation of duties

Project templates

Standardized, version controlled, project templates that separates data, model & feature code, containers, pipeline steps, dags and documentation

Model pipelines

Set of interconnected steps that allow building, training, validating, and registering your models in a standard way

Integrate with existing ecosystem

Integrate AI/ML Platform with company's data repositories, data visualization tools, workflow management platforms…

CI/CD Template

Pre-configured, automated CI/CD pipelines that make it easy for non-technical users to build and deploy the model in a mature and governed way

Model Serving

Each company has specific requirements for model serving that are not easy to implement using cloud-provided capabilities but can still be templatized.

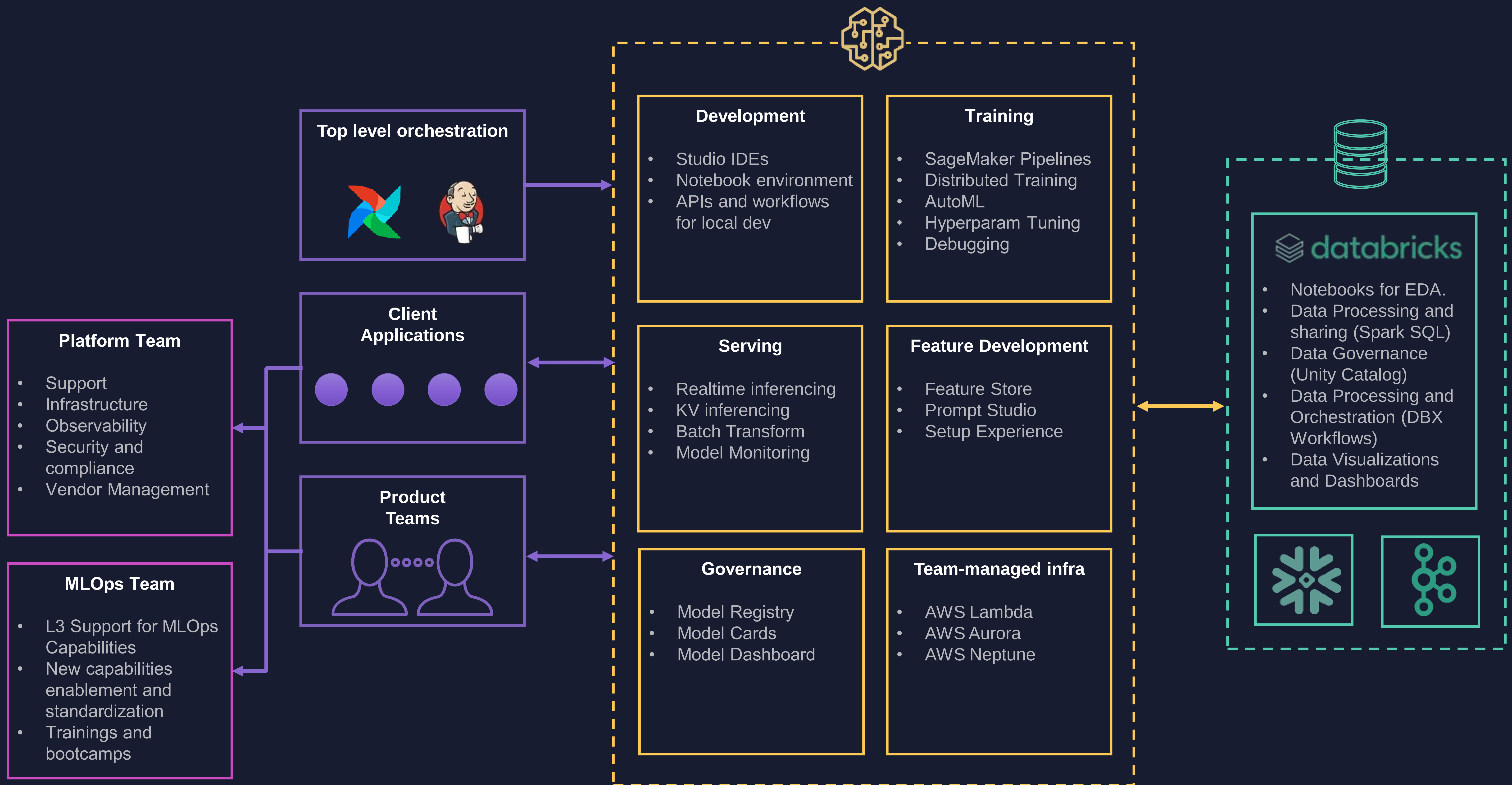
Model monitoring

Centralized monitoring for ML models and products that allows tracking metrics and KPIs that are crucial for a particular enterprise

Migrations

Moving existing ML products and pipelines to the new ML Platform cannot be automated and will always require engineering support.

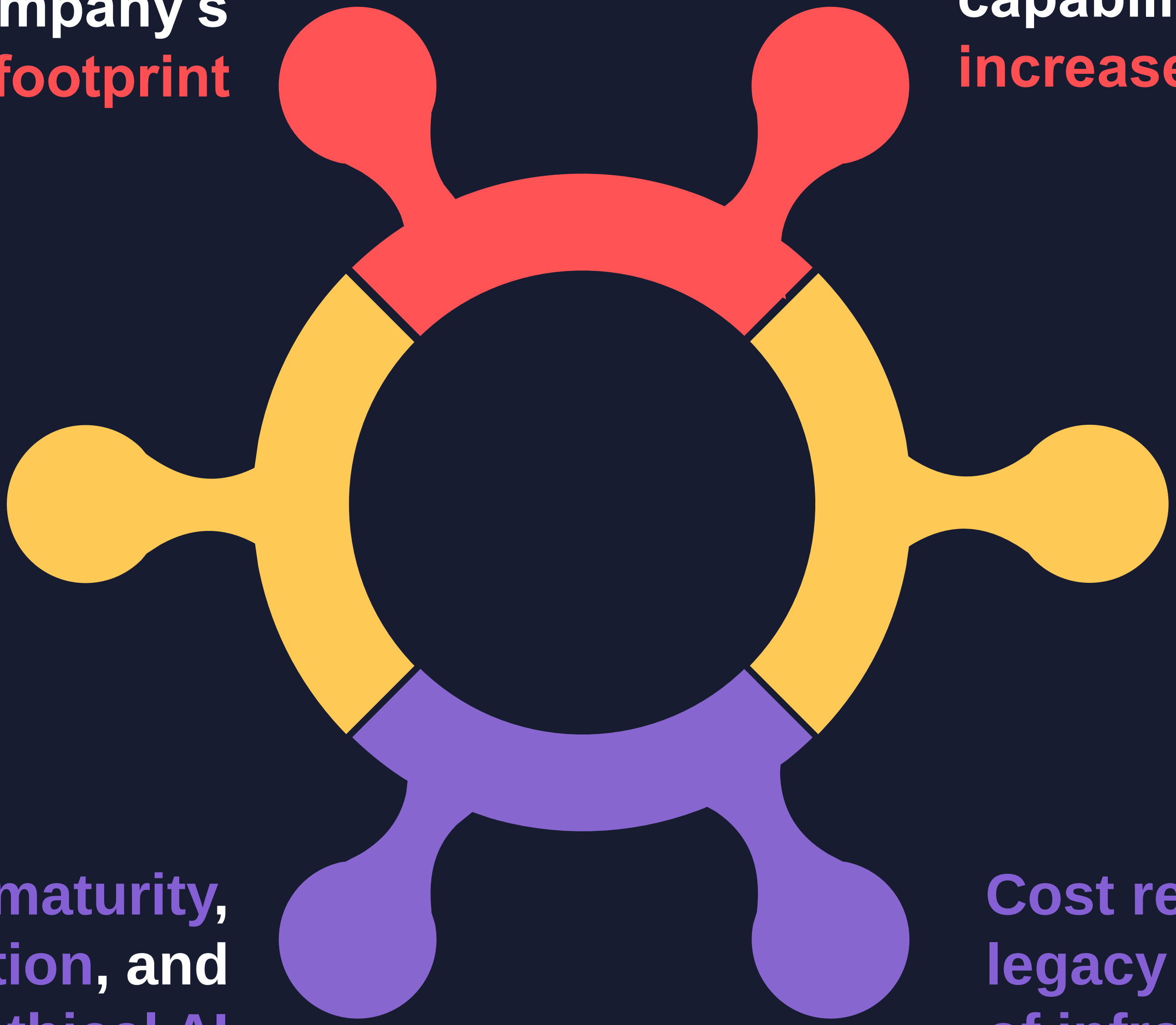
AI/ML Platform Capability Map



What results should you expect?

Standardization and **simplification** of company's technology **footprint**

Enablement of **horizontal scaling** of DS capabilities **without** simultaneous **increases** on the **IT** side



Lowering the entry barrier for new DS initiatives and **democratizing** access to **AI**

Significant **increase in satisfaction** of **users** who create and/or use ML solutions in their daily work

Increased **MLOps maturity**, proactive **risk mitigation**, and **ethical AI**

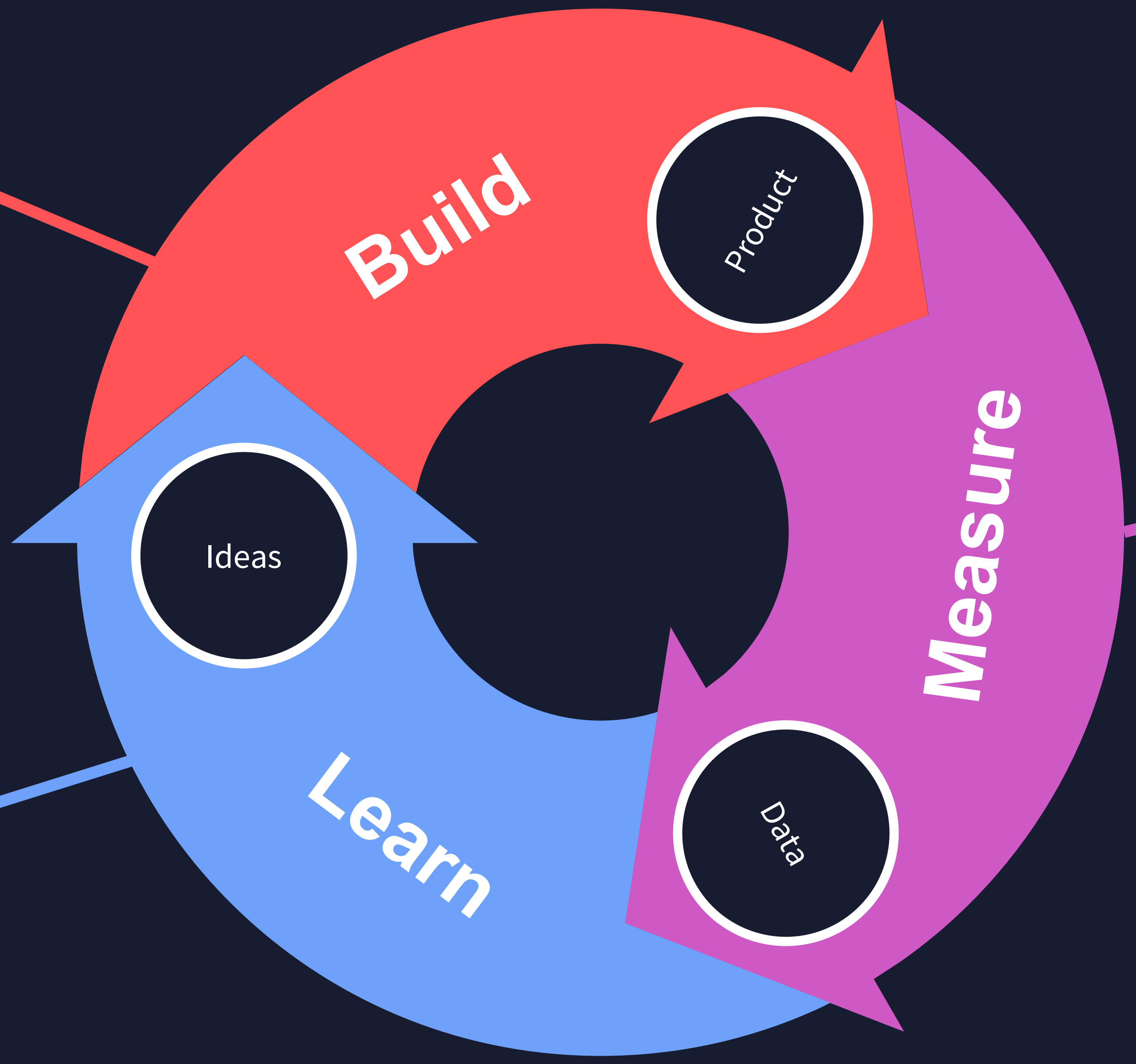
Cost reduction driven by **sunsetting legacy** solutions and **lowering** the **cost of infrastructure**

Lessons learned

The guiding principles worked, but the execution could have been better…

- Use-case driven
- Build once, than standardize
- Small batches

- Analyze the right metrics at the right time
- Act on metrics



Customers
- Platform Awareness
- Platform Capabilities Adoption
- User’s feedback

Products
- Maturity level / Mitigated risks
- Process effectiveness
- Lead time
- Revenue and ROI growth

Platform
- Level of automation
- Capabilities adoption
- Models in production
- Lead time for infra changes

Build-Measure-Learn Feedback Loop
The Lean Startup, Eric Ries, 2011

Thank you!