# MLOps Culture for Continuous Experimentation

**Olga Tsubiks**
DIRECTOR, STRATEGIC ANALYTICS AND DATA SCIENCE
ROYAL BANK OF CANADA

Olga Tsubiks is director of advanced analytics and data science at the Royal Bank of Canada, where she's responsible for the development and evolution of next-generation capacity modeling.

A passionate AI/ML leader, she's spent the last 15 years in various senior roles in data science, big data, data engineering, analytics, and data warehousing.

She's also worked on various data science and analytics challenges with global organizations such as the UN Environment Programme World Conservation Monitoring Centre, the World Resources Institute, and prominent Canadian nonprofits such as War Child Canada and Rainbow Railroad.

**Olga Tsubiks**
Director, Strategic Analytics and Data Science
RBC

# Agenda

Our MLOps discussion will focus on  what happens after your first model is deployed.

- MLOps process used for designing machine learning experiments
- Explain why continuous improvement often requires a cultural change
- Decipher MLOps concepts such as model drift and monitoring (which are useful for ML engineers and can help practitioners who are working closely with data scientists or those who aspire to build complex experimentation frameworks)
- Explore common pitfalls on the way to MLOps maturity

You'll come away knowing how to navigate the challenges of managing ML pipelines and build a culture of continuous experimentation and improvement.

# ML Lifecycle

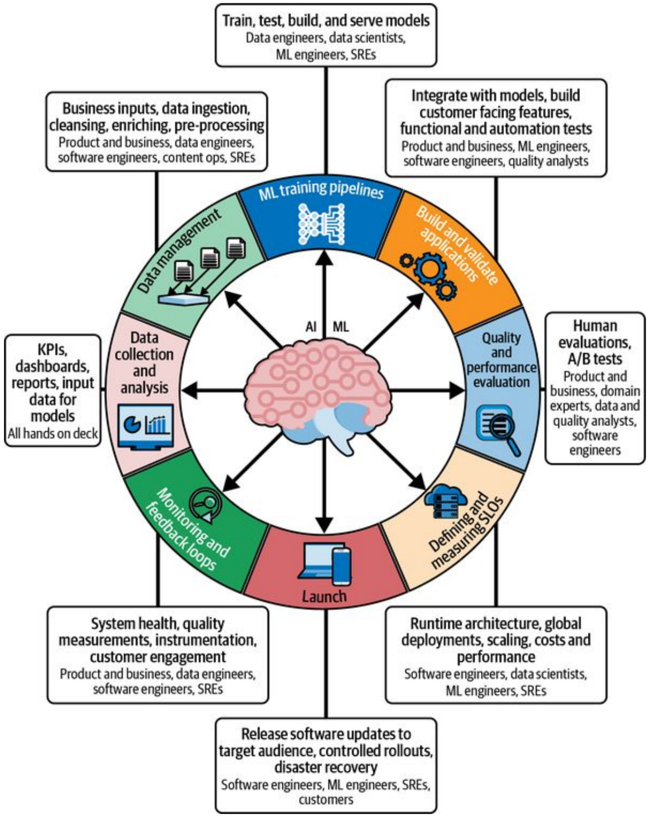



Image Source: Reliable Machine Learning. C. Chen, N. Ri. Murphy, K.Parisa, D. Sculley, T. Underwood. O'Reilly Media, Inc.

AI/ML products are expensive to build, and businesses expect them to be around for a long time.

Experimentation is key to extending the lifetime of the AI/ML product.

# MLOps Maturity and Experimentation

| Level | Description | Technology | Experimentation Highlights |
|---|---|---|---|
| 0 | No MLOps | Manual builds and deployments<br>No centralized tracking of performance<br>Manual model training | No ongoing experiments<br>Manual model improvements mostly to fix errors and defects<br>Infrequent releases<br>Lack of performance monitoring<br>The models fail to adapt to changes in the dynamic environment or changes in the data that describes the environment. |
| 1 | DevOps but no MLOps | Reliance on DevOps team<br>Automated builds and testing of code | |
| 2 | Automated Training | Automated model training and tracking<br>Model management<br>Releases are manual, but low friction | Deployment of the ML training pipeline<br>Manual model improvements |
| 3 | Automated Model Deployment | Automated releases<br>Model traceability and observability<br>The entire environment managed: train > test > production | Integrated A/B testing and model performance and deployment<br>Centralized tracking of model performance |
| 4 | Full MLOps Automated Operations | Full system automation and monitoring<br>Production systems provide information on how to improvise or automatically improve with new models | Automated experimentation for model performance and sophisticated experimentation pipeline that drives the AI product development |

Adopted from Microsoft Machine Learning Operations maturity model

# MLOps and Experimentation: Manual Process

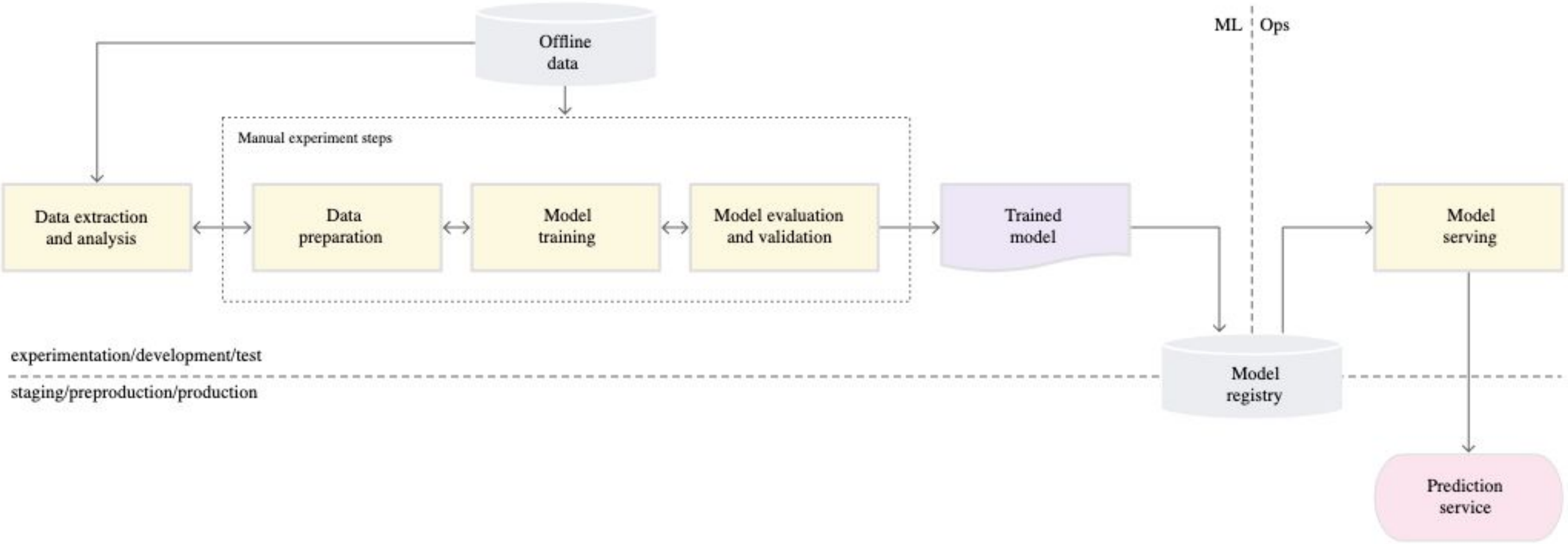



Image Source: Google MLOps Continuous delivery and automation pipelines in ML

# MLOps and Experimentation: Automated Process

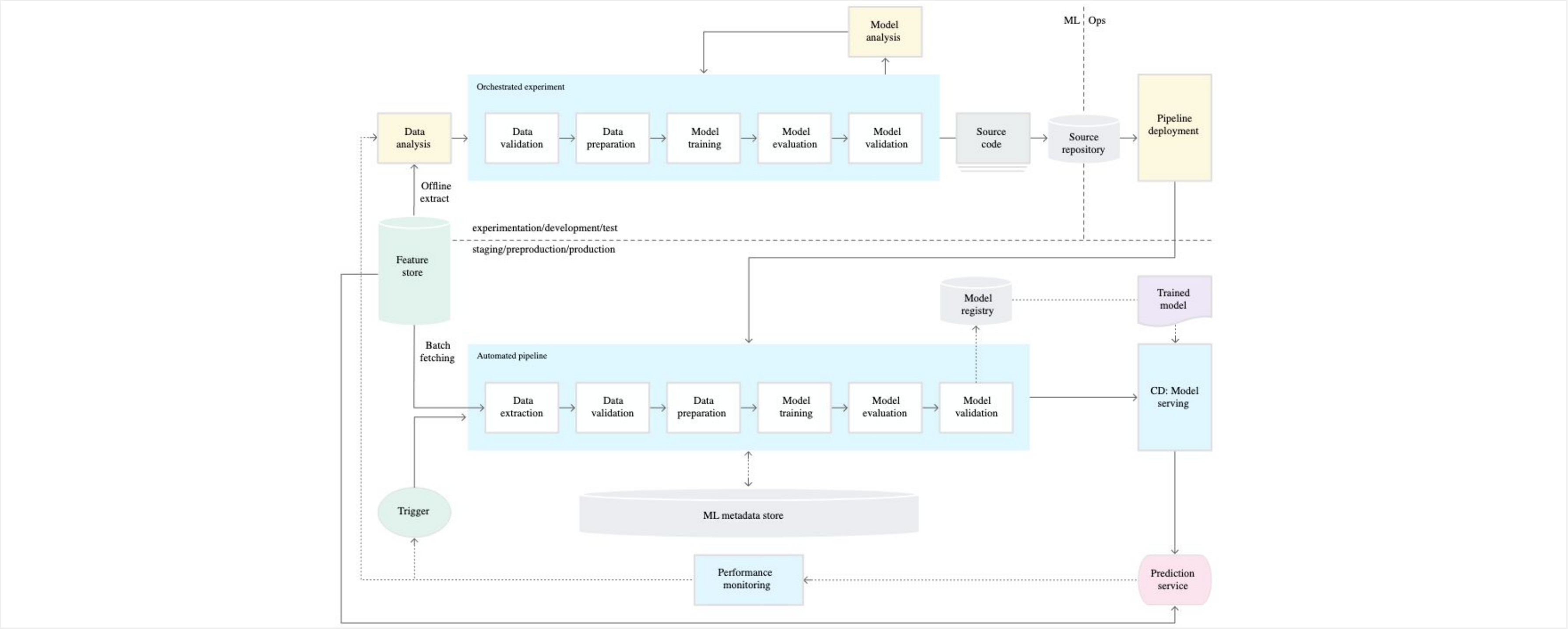



Image Source: Google MLOps Continuous delivery and automation pipelines in ML

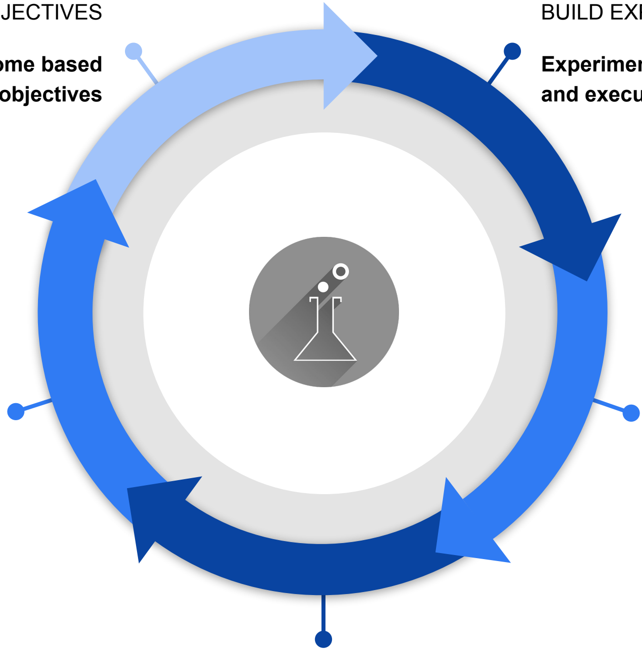
# ML Experiment Lifecycle

SET OBJECTIVES

**Outcome based experiment objectives**

BUILD EXPERIMENTS

**Experiment design set up and execution**



DEFINE PROBLEM(S)

**Driven by organizational strategy, product vision, and model goals**

ANALYZE RESULTS

**Monitoring of experiment outcomes and decisioning**

LAUNCH

**Deployment of experiment and model changes in production**

*Model drift is the degradation of models' predictive power over time as a result of changes in real-world environments.*

*Model drift is our friend. It is a signal that model experimentation is overdue.*

# Culture of Continuous Experimentation

1. Become an advocate of experimentation
2. Lead by example
3. Solicit improvement ideas from others
4. Empower with tools, skills and experimentation frameworks
5. Sustainable experimentation is a set of small, incremental improvements
6. Celebrate experimentation results and communicate outcomes
7. Make your experimentation methodology simple and transparent

# Common pitfalls on the way to MLOps maturity

**Deploying**

ML models are not deployed to production

Lack of standardized deployment process

Overly complex path to prepare models for production

Long backlog for production deployment

**Lifecycle**

Models are not being updated in production

Model drift is not being actioned on

Significant manual work is required to maintain existing models making any new ML model deployments rare

**Monitoring**

No monitoring is performed

Manual monitoring based on data scientist availability

Inconsistent monitoring

No actions are taken on flags raised by monitoring

**Governance**

Lack of access control to input data, model code, and output data

Model results are hard to trace

Lack of data governance to ensure that it is free of bias and complies with regulations

Thank you!