The Data Bistro

A place for business leaders and techies to learn about the latest and greatest data solutions.

 Menu

- Home
- Posts
- About Us

Posted on: April 8, 2021April 8, 2021

Moving Towards Scalable, Responsive, and Sustainable Machine Learning

More organizations than ever are turning to machine learning (ML) and artificial intelligence (AI) to address diverse operational needs.

Traditionally, data modeling projects have employed CRISP-DM (Cross-Industry Standard Process for Data Mining), a blueprint that has been around since the 1990s that employs several iterations of system development and a final deployment. CRISP-DM separates development, deployment, and maintenance across unintegrated teams, which makes incorporating new data or functionality difficult after the initial build period. This is particularly incompatible with a dynamic cloud environment, limiting the integration and responsiveness of the system.

As the landscape of ML and AI expands, work is moving into the cloud-based world of "aaS", which means that instead of selling a static AI model as a product, they provide a dynamic and responsive AI system as a service. This provides the nimble capacity to address business problems not just at one moment in time, but for the foreseeable future.

MLOps (ML + Operations) is an updated blueprint that applies the software development principles of CI/CD (continuous integration/continuous deployment) to analytics projects, including ML and AI. MLOps expands the CI/CD scope to include CT (continuous training) in

order to grow and respond to changing business problems. This provides a constant loop of feedback integration, from development to deployment of analytical components to the system.

In the MLOps blueprint diagram below, automation of routine development tasks, such as data preprocessing and model retraining supports continuous improvement. Augmented intelligence (including augmented analytics) eases the burden on developers to enable more frequent training, validation, and versioning. New data, new types of data, and new requirements continuously feed MLOps development. Performance triggers also spark redevelopment when necessary. Human-in-the-loop workflow provides both the efficiency boost of task automation and higher quality of human validation and verification.
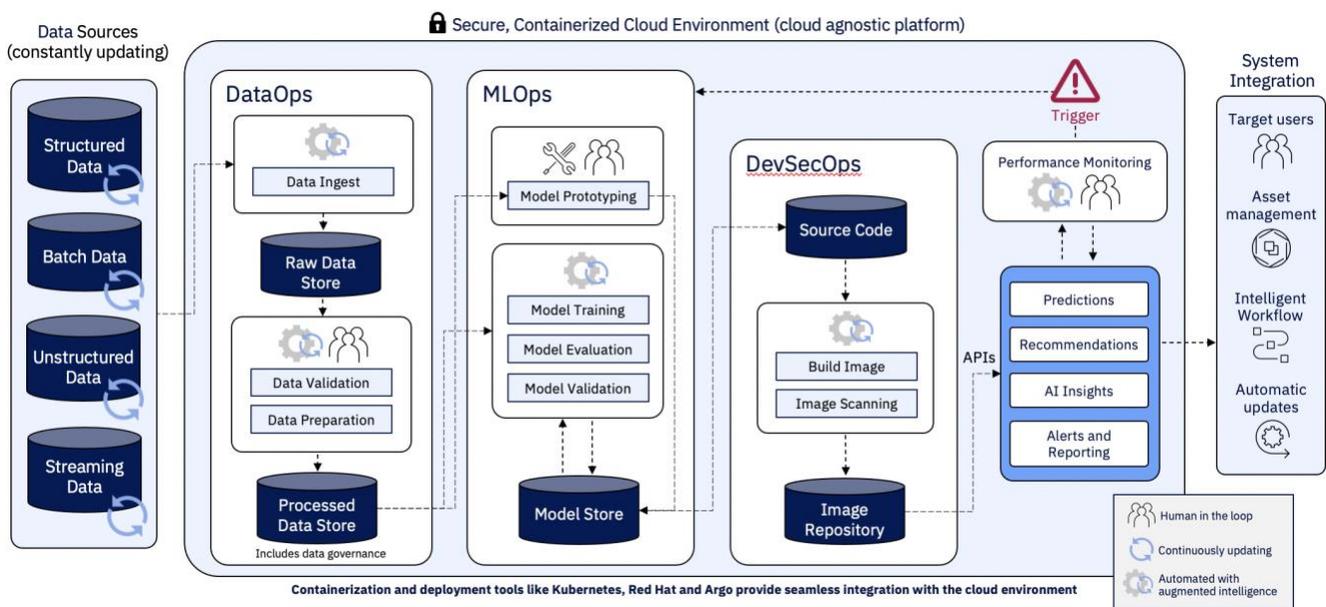
As an example, if a midsize eCommerce company needs predictive ML models to inform targeted advertising, they will likely want models responsive to new shoppers and changing consumer trends — the data on which the models are based should be constantly updating.

- Using CRISP-DM, after the initial models are deployed, each time a new trend emerges a human would have to notice and manually rework and redeploy the models.
- Using MLOps, the system automatically ingests new data and triggers model retraining at regular intervals, seamlessly updating the system.

While CRISP-DM has been the industry standard for decades, enterprise-grade development in today's market requires a more robust framework. MLOps provides an integrated and dynamic approach that can keep up with the latest market demands of ML and AI as a Service.

##

Jordan Fischer is an AI Developer in the U.S. Federal Cognitive and Analytics Group within IBM Global Business Services. She can be reached via email at jordan.j.fischer@ibm.com

---