



















Using CDN:

Some devops team topologies: <https://web.devopstopologies.com/>

Some comparison on ML/AI and Devops processes

|  |  |  |
| --- | --- | --- |
| *Workstream* | *Description* | *Automation* |
| *Data Analysis* | Includes data acquisition and focusing on exploring, profiling, cleaning, and transforming. Also includes enriching, and staging data for modeling. | Develop scripts and tests to move and validate the data. Also create scripts to report on the data quality, changes, volume, and consistencies. |
| *Experimentation* | Includes feature engineering, model fitting, and model evaluation. | Develop scripts, tests, and documentation to reproduce the steps and capture model outputs and performance. |
| *Release Process* | Includes the process for deploying a model and data pipeline into production. | Integrate the AI/ML pipeline into the release process |
| *Operationalization* | Includes capturing operational and performance metrics. | Create operational instrumentation for the AI/ML pipeline. For subsequent model retraining cycles, capture and store model inputs, and outputs. |
| *Model Re-training and Refinement* | Determine a cadence for model re-training. | Instrument the AI/ML pipeline with alerts and notifications to trigger retraining. |
| *Visualization* | Develop an AI/ML dashboard to centralize information and metrics related to the model and data. Include accuracy, operational characteristics, business impact, history, and versions. | n/a |

An automated end-to-end process for the AI/ML pipeline can accelerate development and drive reproducibility, consistency, and efficiency across AI/ML projects.

**Versioning**

Versioning is about keeping track of an application’s artifacts and the changes to the artifacts.

In software development projects this includes code, scripts, documentation, and files. A similar practice is just as important for AI/ML projects because—typically—there are multiple components, each with separate release and versioning cycles. In AI/ML projects, the artifacts could include:

* Data: training data, inference data, data metrics, graphs, plots, data structures, schemas
* Models: trained models, scoring models, A/B testing models
* Model outputs: predictions, model metrics, business metrics
* Algorithms, code, notebooks

Versioning can help provide:

* Traceability for model changes from multiple collaborators
* Audit trails for project artifacts
* Information about which models are called from which applications

A practical example of the importance of versioning for the AI/ML team happens when the performance of a model changes unexpectedly, and the change has nothing to do with the model itself. The ability to easily trace back inputs, dependencies, model, and data set versions could save days or weeks of effort.

At a minimum, decide on a consistent naming convention and use it for the data files, folders, and AI/ML models. Several different teams will be involved in the modeling process and without naming conventions; there will be confusion over which data sets or model versions to use.

Ref: <https://azure.microsoft.com/en-au/blog/getting-ai-ml-and-devops-working-better-together/>