
Despite this financial commitment, data science and ML teams continue to experience issues:

- 33% duplicating their work
- 27.6% rewriting models after team members leave
- 27% justifying the value of their projects to the wider business
- 24.6% slow and unpredictable AI projects

Data scientists today are reliving the chaos software engineers faced in the 1990s.

BACK TO THE FUTURE:

Software Engineers

THEN

Data Scientists & ML Engineers

NOW

Work split across development, testing and operations silos
Lack of version control and continuous integration
Software was delivered in months rather than the minutes it takes nowadays
Difficulties collaborating; work split across various silos

ML and AI projects are more challenging
Requires run tracking which includes version control for training and test data, code and environment, metrics and hyperparameters for each training run
Lack the proper tools required to collaborate on, build and deploy AI models efficiently

Top tools ML engineers use to collaborate with each other:
- 44.5% have a shared spreadsheet for metrics which they update manually
- 38% sit in the same office and work closely together
- 33% use Git for code collaboration only, leaving out data, environment and metrics
- 90% either manually track the model provenance—a complete record of all the steps taken to create an AI model—or do not track provenance at all
- 37% don't track the provenance of their AI models at all

Less than 50% of DMN professionals to develop AI projects manually onto servers or VMs.

Less than 50% don't currently, but would like to deploy their models into production

Lessons learned from DevOps for software engineering need to be applied to ML.

With a DevOps approach, organizations can:
- Eliminate the need for manual tracking of the ML model development lifecycle
- Collaborate effectively and remain in control of their AI models

According to data science and machine learning (ML) professionals at organizations where AI is in production: