Apply Foundational DevOps Principles to Accelerate Data, Analytics and AI Delivery

Published 30 June 2021 - ID G00745612 - 17 min read

By Analyst(s): Afraz Jaffri, Robert Thanaraj

Initiatives: Data and Analytics Strategies; Analytics, BI and Data Science Solutions

Data and analytics leaders must optimize processes for data, analytics and AI delivery as demand surges. DevOps for D&A uses Agile practices to break down functional silos prioritizing process improvements, using value stream mapping to provide continuous value delivery.

Overview

Key Findings

- DevOps for data and analytics (XOps) refers to the disciplines of DataOps, ModelOps and MLOps that are designed to increase collaboration, remove bottlenecks and reduce time to value from concept to delivery of data, AI and analytics.

- Culture change is the main challenge that needs to be tackled in order to sustain XOps using Agile practices across software and application development and data and analytics.

- The task of establishing XOps using Agile practices is made worse for data and analytics leaders due to the multitude of silos created between data engineering, analytics, data science and infrastructure teams.

- Adoption of XOps drives business and IT agility and improves customer value at a time when data and analytics leaders are responding to challenges to deliver clear business value for their organizations.

- The proliferation of “ops” tools for data, AI and analytics diverts the attention of data and analytics leaders away from the primary task of optimizing value delivery and removing process bottlenecks.
Recommendations

Data and analytics leaders responsible for implementing strategies, best practices, techniques and technologies, and vendors supporting data and analytics solutions, should:

- Identify bottlenecks in existing data, AI and analytics processes by engaging in value stream mapping. Remove each bottleneck through experimentation and keep measuring time to value.

- Utilize tooling for automation only where bottlenecks can be alleviated by carefully analyzing continuous integration and continuous deployment requirements across data, AI and analytics value streams.

- Form cross-discipline teams across business and IT, and embed continuous learning by encouraging shared responsibility, experimentation and fast feedback loops.

Strategic Planning Assumption

By 2023, 70% of organizations will use value stream management to improve flow in the DevOps pipeline, leading to faster delivery of customer value.

Introduction

DevOps is a mature discipline in software engineering teams, but has only been sporadically adopted within data and analytics teams. However, the bottlenecks for delivering data, AI and analytics solutions are typically the result of siloed working, lack of collaboration and constant duplication of effort across data, AI and analytics workflows. These challenges are made worse by the multitude of silos created across data engineering, analytics, data science and infrastructure teams (see Figure 1). These are the very things DevOps was created to solve in order to deliver value.
Recognizing the need to unlock these barriers, there has been an emergence of multiple “ops” disciplines (XOps) within data, AI and analytics, including DataOps for data engineering, MLOps for machine learning engineering, and ModelOps for analytical and AI engineering. The hype around these practices is often focused around tooling and software to mirror the common functionality of tools used by DevOps engineers.

However, the fundamental principles of culture, automation, measurement and sharing (CAMS) that need to be in place to use these tools for their intended purposes have not been addressed. This has resulted in frustration from development teams that lack the flexibility to deliver value that an “ops” culture encompasses.
Key to this approach is the handover of shared responsibility to engineering teams, understanding where bottlenecks occur through value stream mapping, bringing tooling for automation on board as collaboration increases and defining measurable metrics to evaluate success.

**Analysis**

Engage in Value Stream Mapping to Identify Bottlenecks in Existing Data, AI and Analytics Processes

The primary reasons for adopting DevOps practices among respondents to the 2020 Achieving Business Agility With Automation, Continuous Quality and DevOps Survey were to increase IT and organizational agility and improve the delivery of customer value. Contrast this to the 31% of respondents to the 2019 CDO survey who said they were effective at consistently producing clear business value for the organization. Taking a value-driven approach is also demonstrated in the way DataOps and MLOps appeared in social media discussions (see the Evidence section and Note 1). DevOps and Agile processes such as value stream mapping are key to solving this challenge for data and analytics leaders.

A value stream map is a visual representation of the end-to-end data and analytics value stream created by all involved parties. Data and analytics leaders should map value streams to identify and eliminate unnecessary activities, and optimize the path from inception to delivery.
One of the main ways in which value can be seen to flow through an organization is through the use of value stream maps to visualize end-to-end software delivery processes. The mapping needs to be carried out by all parties involved in the stream — from the business analyst to database administrators and infrastructure engineers. Their input can be critical in determining what factors are to blame for slow release or cycle times. Value mapping also provides all parties with a view as to what each team's role is in the stream.

In many organizations, data scientists, analysts and data engineers work on a reactive basis, handling tasks that are assigned to them but with no real understanding of how their work fits in the overall process. Changing this approach is where a collaborative culture starts to be built. Even within product or Agile teams, the last mile of production packaging, configuration and release gets hidden from view without taking a DevOps approach.

Figure 2 shows a simplified version of a value stream map that has been created for an application used by customer service team members to help identify the likelihood of customer churn. Each box represents a process that is carried out, with the outputs leading into another process carried out within the same team or a different team. There will be subactivities under each box that can also be depicted in the flow.
Each box should also be tagged with the time it takes to complete the process — the process time. The total time it takes to complete the entire end-to-end flow is the lead time. Measuring these two metrics will reveal that, while in this example the time spent “doing” or adding value was eight weeks in total, the whole flow was completed in six months due to inefficiencies. The goal of the team in subsequent sprints or other application development projects is to reduce the time between the process and lead times, ultimately reducing time to value. This could be done by taking out some of the boxes to create the optimal value flow or by reducing the tasks or duration of tasks that are done within and between processes.
Figure 2 also shows where each XOps discipline fits in the overall flow. When building products, services or applications touch on multiple “ops” disciplines, the risk is for more silos to be created rather than removed. This situation can be avoided by ensuring that every change in an individual workstream is made visible to other workstreams, regularly reviewing the overall process to ensure there is no redundancy and embracing the systems thinking approach of DevOps methodology.

As a team becomes familiar with identifying and reducing the amount of waste in the system (i.e., adopting lean along with Agile), the lessons can be fed back to other teams so that DevOps processes can be adopted at scale. When these initiatives reach a critical mass, it could be the time to introduce a platform team to work on creating infrastructure and processes that reduce waste across all value streams (see Using Platform Ops to Scale and Accelerate DevOps Adoption).

Utilize Tooling for Automation Only Where Bottlenecks Can Be Alleviated

DevOps is forever associated with automation. However, Gartner’s 2020 Achieving Business Agility With Automation, Continuous Quality and DevOps Survey shows that, among survey respondents, automation is used in 56% of DevOps activities (see the Evidence section). Automation is a clear part of any DevOps initiatives but, even if every part of a delivery process could be automated, that doesn't mean the end result will be delivering value. Automation therefore needs to be looked at through the lens of processes that contribute to or diminish value delivery.

Taking a structured approach to automation within DevOps initiatives for data and analytics teams requires analyzing the current value map as shown in the previous section and then ringfencing those activities that can be improved through greater collaboration and sharing, and identifying those where automation can play a part.

Within the cross-disciplinary team, there should be team members who can work on a specific automation goal on each sprint in order to judge the effectiveness against the overall goal of reducing process times and lead time. A set of candidate processes where automation could reduce process time should be marked on the value stream, as shown in Figure 3.
Each XOps discipline has its own tooling that can be used for automating parts of the development process. It is important, however, to not treat the well-known Continuous Integration/Continuous Deployment (CI/CD) objective as the same for each discipline. An application that consists of code only can be version controlled, branched, merged, tested and deployed using automated tooling. Data and models, however, are different, especially when it comes to machine learning. The data, transformation pipeline, features, model, code and results all need to be taken into consideration with tooling still in an emerging state (see Understanding MLOps to Operationalize Machine Learning Projects).

The specific set of automation tools should also be complemented with tools used by DevOps engineers to deploy and maintain applications, servers and infrastructure, including the amount of infrastructure as code, immutable deployments, loose coupling and canary deployments. More details can be found in Gartner’s Market Guide for Infrastructure Automation Tools.
The delivery of automation within XOps disciplines across value streams is the responsibility of a platform team that can unify the tools that are being used across the value stream. There are emerging solutions where all the tools and processes required to create, deploy and manage specific types of workloads can be orchestrated under a single view (see Predicts 2021: Operational AI Infrastructure and Enabling AI Orchestration Platforms).

Form Cross-Discipline Teams Encouraging Shared Responsibility, Experimentation and Fast Feedback Loops

Gartner surveys consistently reveal culture change as one of the top impediments for establishing Agile and DevOps practices among respondents. In addition, Gartner's 2019 Chief Data Officer Survey found 46% of respondents citing culture challenges to accept change as the No. 1 roadblock most critical to data and analytics teams’ success (see the Evidence section).

From a delivery team perspective, creating a culture in which XOps practices can flourish involves three main objectives:

- Engaging team members in the end-to-end value stream, including business needs and outcomes
- Rebalancing team structures to form cross-functional teams, ensuring breadth and depth across disciplines
- Empowering teams to challenge, change and adapt processes as they learn

The set of XOps practices for data and analytics concentrates on specific areas; however, when considering the entire value stream, it is critical to ensure that changes in one piece of the delivery pipeline do not cause problems in downstream or upstream processes. For example, shifting from offline batch scoring to online models that score per request will deliver faster predictions, but may overwhelm application servers.

Creating cross-functional teams (i.e., fusion teams) with data managers (such as data engineers, data architects and data stewards), model developers (such as analysts and data scientists) and software engineers (such as DevOps tooling experts, test engineers and automation engineers) will enable all members of the team to see where bottlenecks are occurring and view the entire backlog of requirements, not just those for their own area.
A people-centric approach is critical to the success of XOps initiatives. Even the best tooling will never make up for shortfalls in collaboration, flexibility, autonomy and openness.

The ability to broaden the scope of expertise of developers so they understand and can even carry out some tasks outside of their core area is referred to as making them T-shaped. This does not mean everyone needs to be a master of every discipline, but everyone should have enough knowledge to recognize how their own design decisions can impact other teams and should be able to communicate these choices to other team members. For example, data scientists should recognize how containerization or code refactoring impacts the choice of libraries used to build models, and operations engineers should recognize the ways in which model drift can occur in order to support model maintenance.

Cross-training and diverse workloads can boost team morale as well as adopting the practice of blameless postmortems, removing the fear of committing mistakes, which in turn encourages learning and development and the ability to fail fast and pivot as needed. When trying out new processes and tools, there should be a recognition that not everything will fully succeed. This requires a shift to managing a portfolio of work where, for some initiatives, the outcome is unpredictable but the value is high. An approach to categorizing AI projects in this way is defined in Gartner's PRISM framework (see Uncovering Artificial Intelligence Business Opportunities in Over 20 Industries and Business Domains).

The time to market of analytics and AI models can be reduced by adapting modular design principles (continuous build-measure-learn feature cycles). Business value can also be delivered more frequently by limiting the feature scope and data scope.

Finally, shift responsibility for how products are delivered from individuals to teams, and celebrate successes and investigate failures at a team level. Draft the responsibilities for the team, not for individual roles. Team success and failure should be measured using two categories of metrics:

- Business impact metrics, such as customer satisfaction, sales value, subscription revenue or market share
Progress metrics, such as mean time to value, story points, code quality (bugs in production) and reduction of technical debt

A complete hierarchy of metrics used to measure Agile delivery is described in Choose the Right Metrics to Drive Agile, DevOps and Continuous Delivery.

Data and analytics leaders must focus on these foundational principles of DevOps and create an environment where they can be tested, broken, rebuilt and reused to reduce time to value of data and analytics delivery. A detailed description of the many ways in which culture change can be achieved is described in Data and Analytics Leaders: Rewire Your Culture for an AI-Augmented Future.

Combining tooling with a change in culture, communication and ways of working will lead to the successful establishment of DevOps initiatives within data and analytics teams. The need for data and analytics leaders to start their journey has never been more pressing, but there are many hard-learned lessons from software engineering that can be used as a guide. Focusing on these principles can bring transformational change in the way data and analytics is used to create value propositions for organizations.

Evidence

Achieve Business Agility With Automation, Continuous Quality and DevOps Survey: In answer to the question, "What is your best guess of the proportion of your organization’s DevOps activities that is currently manual versus automated?" 56% of DevOps activities were automated, on average.

This study was conducted online from June 2020 through August 2020 among 205 respondents working for service providers, cloud providers and end-user organizations in North America and Western Europe that have deployed or are using DevOps.

Qualified organizations had at least $500 million in annual revenue and were required to primarily operate in the banking and financial services, government, insurance, healthcare provider, and retail industries.

Respondents were required to work in their organization’s IT function, have a job title less senior than C-level, and be two or more layers away from the most senior executive in their organization.
Respondent’s role had to be primarily focused on application development, infrastructure and operations, or business intelligence and information management. In these focus areas, they were also required to perform relevant roles/activities.

The study was developed collaboratively by Gartner analysts who follow digital business trends and Gartner’s Research Data and Analytics team.

Results of this study do not represent global findings or the market as a whole but reflect sentiment of the respondents and companies surveyed.

**Chief Data Officer (CDO) Survey:** When asked, “Which of the following are the most important roadblocks to the success of your data and analytics team?” 46% cited culture challenges to accept change.

This study was conducted to explore the business impact of the CDO role and/or the office of the CDO. The research was conducted online from September 2019 through November 2019 among 293 respondents from across the world.

Respondents were required to have the title of CDO or chief analytics officer (CAO), or to have the responsibilities of an executive-level data and analytics leader in their organization (in the case of organizations without an official C-level data and analytics title). The survey sample was gleaned from a variety of sources (including LinkedIn), with the greatest number coming from a Gartner-curated list of over 2,000 CDOs and other high-level data and analytics leaders.

The study was developed collaboratively by Gartner’s Research Data and Analytics team.

**Agile in the Enterprise Survey:** This survey asked participants to identify the top three challenges or obstacles their organization has faced in adopting Agile development. Changing the development culture was the No. 1 challenge.

This study was conducted via an online survey from 3 June 2019 through 25 June 2019 with 130 members of Gartner’s Research Circle — a Gartner-managed panel of IT and IT-business professionals.

Qualified participants included business end users with either an IT or IT-business focus as a primary role. Eighty-seven percent of participants use Agile for at least some of their application development.
The survey was developed collaboratively by a team of Gartner analysts and was reviewed, tested, and administered by Gartner's Research Data and Analytics team in collaboration with Gartner analysts.

For additional queries on the data or survey instrument, please contact anna.toncheva@gartner.com

**ITL Circle Omnibus Survey**: Gartner's ITL Circle 2019 Omnibus Hot Topics Survey showed organizational culture as the No. 2 ranked barrier when participants were asked to identify the top three barriers to success on their organization's DevOps journey.

This study was conducted via online global omnibus from 23 October 2019 through 7 November 2019 with 110 IT leader members of Gartner's Research Circle — a Gartner-managed panel.

Forty-nine IT leaders qualified for and completed the DevOps section intended to explore current adoption and implementation challenges to DevOps. Questions were designed for responses from IT leaders who are familiar with DevOps adoption at the organization.

Respondents were primarily split between the U.S. and Western Europe. The survey was developed collaboratively by Gartner analysts focused on I&O and Gartner's Research Data and Analytics team.

**Social Media Analysis**: DataOps mentions have remained steady compared to mentions of MLOps that have seen huge increases over the past two years. Discussions mentioned that this is due to data-led operations in organizations shifting from data quality through automation and continuous integration to context-driven data quality in AI and ML models incorporated under the MLOps discipline. This is also because AI-driven automation, governance and data protection are becoming integral parts of business strategies.

**Note 1: Methodology**

**Approved Methodology**: Gartner conducts social listening analysis by leveraging third-party data tools to complement or supplement the other fact bases presented in this research. Due to its qualitative and organic nature, the results should not be used separately from the rest of this research. No conclusions should be drawn from this data alone. Social media data referenced is from 1 January 2019 through 19 September 2020 in all geographies and recognized languages. However, it is important to note that the data for China is largely nonrepresentative due to the restrictions that China places on foreign-owned social media platforms.
Sources Covered: By default, social media sources considered for analysis include Twitter, Facebook (publicly available information only), aggregator websites, blogs, news, mainstream media, forums and videos (comments only), unless and until specified.

The SMA Team: Fahim Talmeez from the Social Media Analytics team contributed to this research.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Use 8 Simple Steps to Get DevOps Right
Demystifying XOps: DataOps, MLOps, ModelOps, AIOps and Platform Ops for AI
Top Trends in Data and Analytics for 2021: XOps
Introducing DataOps Into Your Data Management Discipline
Client Question Video: How Do We Sell the Business Benefits of Data and Analytics?
How to Build a Data Engineering Practice That Delivers Great Consumer Experiences
Operational AI Requires Data Engineering, DataOps and Data-AI Role Alignment