

WHITE PAPER

The Cultural Impact of DataOps: Collaboration, Automation and the War on Silos

By Hitachi Vantara
September 2019

Contents

Introduction	3
Collaboration	4
Automation and Metadata Mindset	4
Understanding Data as a Shared Asset	6
End-to-End Design Thinking	6
Enlightened and Guided Empowerment	6
Silo Paranoia	7
Push Down Decision-Making	7

The Cultural Impact of DataOps: Collaboration, Automation and the War on Silos

Science fiction author William Gibson captured a powerful pattern about how knowledge, expertise and technology are disseminated when he said, “The future is already here — it is just not very evenly distributed.” This quote applies to many realms of technology but it has special resonance and explanatory power in the emerging practice of DataOps.

Hitachi Vantara believes that DataOps represents the future of data management. DataOps, an evolution of the practices of DevOps applied to enterprise data challenges, solves many of the problems companies now face in unlocking the power of data by expanding its use in every corner of an organization.

Two companion papers lay out our vision for DataOps and how it will be implemented. In this paper we cover the cultural changes that DataOps will bring and that’s where William Gibson’s quote becomes so instructive.

The first crucial idea is that, like the future, DataOps is already here, concentrated in certain industries and organizations and is just starting to be more evenly distributed. If we look at data-rich web-scale companies, like cutting-edge hedge funds or ad tech firms that run millions of auctions a minute, we can see implementations of many DataOps practices.

These types of companies have the ability to build and maintain numerous data pipelines that deliver data to dashboards, applications, autonomous systems and anywhere else it is needed. These companies are leading the way into the DataOps future and the rest of us need to do something similar, in a way that works for our organizations.

But as DataOps becomes more broadly adopted, it is clear that the practice must be tailored to the unique needs of an organization. At Hitachi Vantara we believe DataOps must evolve to address issues such as data governance, security and access control in an automated, scalable fashion. The early adopters of DataOps have not solved these issues because while they operate with huge data volumes at high speeds, the organizations are much simpler and smaller. In addition, the underlying infrastructure supporting DataOps must be built to evolve to achieve operational agility. Again, at the early adopters, agility is achieved through a strong DataOps culture and by throwing huge engineering teams to build custom systems to achieve needed levels of automation in both applications and infrastructure. See our companion papers for a detailed discussion of how DataOps will work for the enterprise.

But part of the evolution will be cultural and that’s what we want to talk about in this paper. We see the following cultural dimensions as crucial to successful adoption of DataOps:

- ▶ Collaboration
- ▶ Automation and a metadata mindset
- ▶ Data as a shared asset
- ▶ End-to-end design thinking
- ▶ Enlightened and guided empowerment
- ▶ Silo paranoia
- ▶ Push down decision-making

Many of these areas overlap and support each other. The end result for DataOps, as it was for DevOps, will be an organization that is more unified and works together at a much faster clock speed.

Collaboration

Collaboration in the world of DevOps is simpler than it will be in DataOps. DevOps unifies the worlds of IT operations and software development. These two powerful and complex functions were relatively focused and highly technology oriented.

In DataOps, the landscape is much bigger. You're not just talking about two engineering-oriented disciplines. DataOps encompasses everyone from the beginning of the data supply chain where data originates (from IoT devices to enterprise applications to massive third-party repositories such as the Open Data Initiative) to all the people who model and blend data all the way to those who put it to use in applications and analytics.

Collaboration in DataOps thus has many more dimensions than it does in DevOps. In the real world of day-to-day business, where end users put data to use, they must be able to declare their problems and collaborate to pull applicable data together on their own while at the same time feeling free to ask relevant specialists for help, from data engineers to data quality experts to data scientists. This model of allowing people to use policy-based systems that provide more self-service shouldn't hinder collaboration; instead it should foster collaboration on more meaningful problems and reduce grunt work across the board.

Automation and Metadata Mindset

In the world today, we have people who throw up their hands and solve their problems in one-off ways using isolated spreadsheets or data marts that they maintain themselves. Just getting the job done with spreadsheets has created problems we're all familiar with, including stale data and errors in formulas or links. The world of DataOps embraces automation, with metadata as its foundation to facilitate broad use of the latest data, eliminating the frustration and delays that led people to use spreadsheets. Spreadsheets will remain vitally important but they must be used not in isolation but as part of a "spreadmart" where usage of data is integrated with the rest of the data pipeline.

In DevOps this meant companies had an integrated tool chain, with everything from automated testing, automated assembly, continuous integration/continuous deployment and the ability to essentially push one button and go from source code to deploying an application or website. As a result, change became much safer and frequent iterations became the norm.

With DataOps, users are going to want the same type of operational simplicity and rapid iteration, but to achieve it, all of the functions will have to be automated. That automation will largely be driven by metadata descriptions, including metadata descriptions of data, of processes used to transform data and of repositories and the infrastructure. The automation will be services that sit on that that are either configured by policies or by other metadata.

Once problems are well understood, this type of automation mindset will lead to the kinds of autonomous AI and machine-learning systems we are seeing in a number of areas where systems are able to make decisions that people used to make. This is related to another cultural change that we'll discuss later called push-down decision-making. It is also the way that the larger trend of the cognitive revolution is being implemented. We'll thus start with automation and the expansion of it and then progress to the point where automation is doing more significant work.



Understanding Data as a Shared Asset

Everyone talks about data, but its economic value is less frequently explored (see, for example, Hitachi Vantara's [recent joint research project](#) conducted with the University of San Francisco).

Unlike other assets, data does not depreciate, becoming less valuable the more it is used. Instead, data gains value as it is used. Data is not the new oil; it's far more like sunlight. Its power can radiate and shed light on different aspects of a business all day long, all week long, all year long.

In the past, organizational boundaries created ownership around certain types of data. In order to access that data, you would have to go through the organization that owned that data.

The instinct has been to control data and this won't go away easily, as data is a valuable resource and therefore a source of power. But we need to have the governance, access control and security to enable prudent sharing, because unlimited sharing doesn't work in most venues.

If companies allow hoarding of data, it can lead to new silos (which will be discussed later). Assuming control of assets and hoarding data is a long-standing practice. For DataOps, this must change. Sharing must be the baseline assumption.

End-to-End Design Thinking

One way to understand both DevOps and DataOps is as an application of design thinking. In both DevOps and DataOps the whole problem, from end-to-end, including all the goals, is rethought. The scope of the problem is much wider than it was in the past. Many of the principles of design thinking are in play in DataOps. DevOps expanded the scope of the problem, seeing it not as a Dev problem or an Ops problem, but a DevOps problem, which included the impact on the users.

DataOps does the same thing with organizations thinking through the flow of data from its creation to its use. But with DataOps, far more groups are impacted, as the entire organization relies on data. The key transformation is not that specialities won't exist anymore but putting everything in the service of a larger automated system allows people to create and maintain data supply chains that operate efficiently.

End-to-end design thinking can be seen in DataOps and it can also be seen in many of the social innovation practices where governments and other organizations work at solving larger problems. DataOps can be considered a sort of social innovation in the enterprise.

Enlightened and Guided Empowerment

DataOps is really about getting more people involved and allowing them to do more work supported by automation and intelligent systems. That, of course, can be described as self-service. But if we think of self-service only in the way that the term has been used in the past, that won't work for DataOps. What we're looking to achieve is guided and enlightened empowerment. We want to provide a simpler way for end users to get the data they need. We want data pipelines to be easier to manage and security and governance less cumbersome to administer. It means making complex technology used for infrastructure easier to manage for operational agility. It means expanding automation at all levels.

Therefore, self-service must not be about just giving someone a simplified system, but one that offers suggestions and guidance. In addition, the idea that just because a person is trying to do this on his or her own doesn't mean he or she will not benefit from the help of others. Self-service shouldn't be a barrier to collaboration with experts. We want data experts to provide help when it is needed. Enlightened and guided empowerment must become a cultural value for DataOps to succeed.

Silo Paranoia

DataOps and DevOps sought to break down silos that emerged naturally between parts of the organization. The impulses that created these silos will not disappear overnight, whether it's attempting to control assets or gain power. Part of DataOps culture is staying vigilant and on the lookout for any time silos begin to reemerge. In a DataOps culture, it should be possible to complain when unreasonable restrictions on using data are stopping productive work and have that complaint heard.

Push Down Decision-Making

It's all well and good to think of data flowing everywhere in the organization, but the impact of that greater spread of data must be considered. If data is flowing everywhere, it should make everyone smarter. The question thus arises: Should users be required to get approval for every decision they make? Therefore, perhaps the biggest cultural change from DataOps relates to decision-making. If data is flowing everywhere and informing more people, it is only logical that they make decisions based on that data.

Push down decision-making, where those using the data make the decisions, is a major cultural change. Too often, people at the edge of the business are not accustomed to making decisions confidently because they have never done so before and often weren't hired to do so. In addition, autonomous systems will play a role in pushing decision-making down and toward the edge. This type of change, where decision-making moves to the edge with more empowered people using data-based insights, alters the approach to management and oversight and builds decision-making "muscle" throughout the enterprise. Companies must be prepared for this cultural evolution and have a plan to handle it for DataOps to be successful. U.S. General Stan McChrystal started the ball rolling in this direction in 2003 with his "team of teams" concept, in which the military ended siloed thinking and enabled the innovations of small teams to be spread throughout the entire organization.

Culture is a broad and tricky subject. This discussion is not intended as the last word on DataOps culture but rather as an invitation to extend the discussion about how DataOps will impact culture at your organization. Hitachi Vantara will be keeping an eye on what we learn internally about DataOps culture and what our customers are learning. We hope you can join us in that discussion. Please take a look at our [blog](#).

Hitachi Vantara



Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
HitachiVantara.com | community.HitachiVantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
HitachiVantara.com/contact