The Chief Data Officer's Guide to Generating Impact

Chief data officers face a daunting challenge: Bring about data-powered business transformation, and do it fast.

Here's a guide to accelerating time to value without spending a fortune — and while generating business impact.

Mark Van de Wiel

Fivetran

For chief data officers, the pressure is on

Confronted by outdated infrastructure and sometimes stymied by naysayers within their own organizations, chief data officers often struggle to bring about data-powered business transformation quickly. The inability to meet often-unrealistic expectations has led to rapid and costly turnover among CDOs.

A recent Harvard Business Review article notes that the average tenure of a CDO is less than 2.5 years. The authors found that the honeymoon for CDOs often ends after 18 months, when boards and fellow executives become disenchanted with the CDO's failure to achieve major change. That's a quick timeline given that a data-focused overhaul of a large company with a lot of legacy infrastructure typically requires a multi-year effort.

How could this happen? In recent years, a legion of marketers trumpeted the power of data to rewrite the rules of business. "Data is the new oil," they claimed. At the same time, companies across the globe began investing billions to bolster their data analytics capabilities.

Lost in all the hype is the fact that becoming a data-driven organization is not something that just happens without a significant effort and investment, especially for companies attempting it on their own.

The challenge for CDOs is clear: How can they use data to create transformational change within just 18 months?

It might seem like a daunting task — and it is. But it's not a hopeless one. In this ebook, we'll detail the challenges data leaders face and then explore how they can accelerate time to value, not by ripping out their existing data infrastructure and replacing it with something else, but by implementing a modern data stack.

A modern data stack is a cloud-based data and analytics environment built around the best-of-breed technologies. It provides a radically new approach to data integration that saves time, money and effort, ultimately allowing engineers and analysts to pursue higher-value activities such as data science and machine learning projects.

A successful data-driven transformation demands a process of augmentation, not reconstruction: Adding automation and connectivity that enables the rapid generation of meaningful analytics from a wide variety of existing data sources.

In this book, we'll show you how to do just that.

What is holding up data-driven transformation?

For an impartial view of how companies are faring in their efforts to become more data-driven, let's take a look at a survey published last year by NewVantage Partners, a management consulting firm.

The much-circulated report showed that though more companies than ever invest in big data, a large number reported being disappointed with the return on those investments.

NewVantage found:

- Only one third of respondents had created a well-articulated data strategy
- Less than half reported driving innovation with data
- Only about 4 in 10 are actually managing data as a business asset

The plodding pace of change isn't what anyone expected with the onset of the data era. The triumvirate of cloud computing, AI and data analytics was supposed to usher in a new order, resulting in radical innovation, substantial cost savings and finally increased profits.

That promise is still good, and the companies that can leverage data successfully are seeing outsized results. But too many companies still can't wrap their arms around the problem.

Why does generating a data-driven business transformation take so long? And why has it proven so challenging?

Numerous journalists, analysts and insiders have weighed in with opinions on what they believed is hobbling transformation efforts. Many of them cite a lack of data literacy, not enough buy-in from important decision makers, a poorly defined CDO role and outdated legacy data environments.

Some point fingers at poor data quality and complicated or ineffective analytics tools. Supporting that last assertion is a **study produced by Deloitte**, which found that 67 percent of C-suite executives and managers said they were uncomfortable accessing their company's analytics tools.

All of these are important contributing factors, but the fundamental issue underlying all of them is a lack of easy access to data. That's a problem that has technological, process, training and cultural dimensions.

Let's take a look at those three aspects.

Elements of a successful transformation

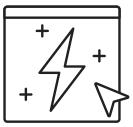
To enable easy access to data and complete the transformation into a data-driven organization successfully, three elements need to come together:

People

Data-driven cultures start at the top, where leaders and managers set an expectation that decisions must be anchored in data. If the organization lacks a data culture, then when someone in marketing or sales has a hunch, the most sophisticated tools will sit idle. Also, if people don't get the proper training in how to use these tools, they won't get the most out of them.

Processes

This is about democratizing access to data within enterprises. Along the same lines, the right tools must be skillfully incorporated into a company's business processes. If solutions are overly complex or demand too much effort to fit into the workflow, then without a doubt they'll go unused.



Technology

For data to be of use to the people on the front lines, they need analytics tools to be intuitive. Such tools must enable uncomplicated access to reliable, easily understood and actionable information.

Achieving this kind of integration isn't easy. The task requires data teams to collaborate and obtain support from a company's business departments so that they can zero in on solving specific, real-world problems. From what we and others have seen, this kind of collaboration at many companies is too often missing. Maybe that's one reason why a **TechRepublic survey** from last year found that 27 percent of business professionals said that "a sharper focus on business-use cases would build value" for the company's analytics.

Another important aspect is ensuring accuracy of data. Eliminating redundancies and normalizing data structures across different sources is vital and must be a continuous process.

In the end, the insights produced are only as good as the available data. A lack of validated information only causes delays, or worse. Poor quality data can cast doubt on results.

Additionally, companies wishing to incorporate data management correctly in the current environment of threat actors and cyberattacks mustn't skimp on security. Leaders should seek out trusted partners and vendors and ensure that they're adopting the highest level of best practices. Those include encrypting data that sits in cloud storage as well as during transit, implementing solid access management systems and asking the cloud provider for details about security.

Cultural issues and how to handle them

Demonstrating a data program is effective is key to winning over important leaders and the rank-and-file. That's because of all the hurdles that CDOs face in their transformation journey, none is more daunting than those that stem from cultural issues.

The NewVantage survey showed that an overwhelming majority (92.2 percent) of the companies surveyed pointed at their company's culture as one of the roadblocks to changing the way companies manage data. That large percentage also represented a significant increase from the survey response of four years ago, when 80.9 percent of respondents cited culture as a significant obstacle.

These cultural challenges can take many forms. What we've seen ties back to those overly optimistic expectations we mentioned earlier. First, depending on the pace at which an organization's culture allows a CDO to move, 18 months is unlikely to allow enough time to implement an effective organization-wide data strategy. Many times, what CDOs end up bumping into is that their organization just isn't ready.

This can manifest itself as a budget figure that's fluid. Initially, a CDO may be promised access to a sum of money and told leadership believes there's an opportunity in data transformation.

Later, after the quarterly or annual results get published and results fail to meet the board's expectations, the trouble begins. The promised financial support becomes a subject of debate on the board or the leadership team. Some begin to argue that it's time to pump the brakes and limit the scope of the data initiative.

Confronted by all this, a CDO may be tempted to swing big and do too much. Attempting to boil the ocean with so little time to deliver is too risky if the strategy gets held up.

A better way to begin is for CDOs to target and achieve critical but more modest goals. A CDO could identify several areas where the organization struggles that can be addressed with data, but also solved relatively quickly with the available resources.

For example, the company may have found that a key variable to signing up new customers is the number of visitors to its website. Let's assume this information was hard to find. The CDO could score a big win by making a dashboard available to executives — or even one that's accessible to the entire company. The dashboard could reveal the number of unique daily site visitors and perhaps even the conversion rate (the percentage of visitors who convert to customers, or to the next step in the sales cycle).

These are the kinds of important, incremental triumphs that will win a CDO trust when he or she first starts. Don't make it more complicated than it needs to be. The point is to create and demonstrate success, champion that success, celebrate it, and then grow from there.

Building relationships is crucial

A CDO who doesn't prioritize the changes required to establish a strong data strategy early on will fail to notch these early wins. That can mean they won't receive the necessary resources or support from department managers for subsequent projects. The CDO may not get access to skilled and experienced engineers or data scientists.

Everything grinds to a halt — and that's when many CDOs start polishing the resume. It's hard if not impossible for them to notch any victories if they run into an unwillingness within the company to embrace the changes required to establish a strong data strategy. Skeptical department managers can sometimes get in the way and data tools go unused or valuable collaborative efforts stall.

Rarely is this caused by malevolence. The problem can more often be traced to inertia, an unwillingness within the ranks to embrace change.

That's why building strong relationships with management early on and winning the trust of all the stakeholders is essential. And many companies are beginning to realize that this plays a key role. A CDO who understands a company's culture, someone possibly who comes from within the organization, stands a better chance of pulling off the transformation.

The power of data in the cloud

Gathering information and poring over the data in search of insights didn't begin in the cloud era or even with the rise of the internet. For centuries, merchants have tried to understand the needs and wants of their customers and the efficiency of their employees. But even with

primitive technologies, business data analysis provided some level of insight that led to improvements in business processes.

For example, in the post-war automotive sector, a group of quantitative analysts used data to completely overhaul production at Ford Motor Company to create greater efficiencies. Later, the emergence of credit and the rise of credit cards were enabled by tracking consumer behavior and calculating risk.

Data analysis even brought fundamental change to Major League Baseball, when a handful of statisticians analyzed the way players were evaluated and discovered the methods were flawed, as dramatized in the film "Moneyball."

Data analysis is essential to business and has been for a long time. What's different now is that the combination of cloud computing and readily available machine learning algorithms has supercharged the ability of companies to track, store, analyze and distribute data and the insights derived from them.

It's almost impossible to overstate the importance of the cloud to the ongoing data revolution.

The cloud is an environment that provides infinite scalability and presents very limited risk. Enterprises can try out technologies without having to plunk down big upfront investments. That makes it easier to get started and scale up from there: Customers pay as they go and only pay for resources they consume.

In addition to cloud-based databases, there are technologies such as Fivetran that help companies gather data and consolidate that data into the cloud environment. Cloud-based business intelligence and data analytics platforms make the data accessible and usable to end users with minimal effort.

With the aid of these new technologies, businesses are achieving astonishing success.

For example, a large US-based company in the entertainment industry, with holdings in the dining, gambling and sports betting sectors, is consolidating data from different operational systems into a cloud environment. This enables managers to pinpoint their target marketing based on all the different ways customers interact with the company's brand. The company also relies on these tools to evaluate the performance of staff and different business units.

An airline benefits by consolidating data and then reaping the related and relevant information its systems provide. For example, each time a dispatcher must decide to reroute or cancel a flight, the technology informs managers of the status of the flight, the number of passengers affected, the weather, etc. The airline is also enabled to perform predictive analysis on maintenance to minimize operational impact and reduce costs.

The dangers of manual pipeline construction

Many data leaders run into trouble when they direct their teams to build custom data pipelines. Fivetran and Wakefield Research **surveyed 300 data and analytics leaders** last year and found that data engineers spend 44 percent of their time on building or maintaining data pipelines.

According to those surveyed, a median of 12 data engineers were involved and the total average annual cost per company was $520,000.

It's worth noting that these maintenance chores aren't value-generating. The engineers simply transfer data from different sources into a warehouse and adjust the pipeline when APIs change or new data attributes become available.

Among the other eye-opening findings in Wakefield's report were the fact that 71 percent of respondents said end users were making business decisions with old or error-prone data, and that 85 percent said that bad data had led their enterprises to make poor decisions that cost them revenue.

Rather than spend time manually building and maintaining data pipelines, many companies have turned to automated, managed solutions, freeing data engineers and analysts to do more productive, value-generating work.

An architecture for success

Regardless of business sector or region, companies have begun to manage at least some of their data systems within cloud services, the largest of which include Amazon Web Services, Google Cloud Platform, Microsoft Azure, Snowflake and Databricks.

The cloud represents an unprecedented opportunity. These cloud platforms include enormously capable, scalable data warehouses, such as Snowflake, Amazon Redshift, Google BigQuery and Azure Synapse, and data lakes including Amazon S3, Azure Data Lake Store, Google Cloud Storage and Databricks' Delta Lake. They enable companies to create and operate data-centric applications and provide sheer infinite scalability to handle enormous amounts of data from a wide variety of sources.

These platforms can help clients leverage powerful analytics tools by giving them a standard underlying data store to work on based on a pay-for-use model.

In order to enable this kind of centralized analysis, many companies have utilized a **modern data stack (MDS) architecture**, which simplifies movement of data from applications and legacy databases into that centralized data store. The elements of a modern data stack include:

- A data pipeline or data integration solution, also known as an Extract, Load and Transform (**ELT**)
- A data warehouse, data lake or lakehouse, where data is organized in a columnar format (to support fast and easy extraction of aggregate data, rather than to support record-by-record transactions)
- A data transformation technology, which converts data from one format into another in order to prepare it for analysis
- A business intelligence or data visualization platform, which enables the creation of charts, reports and dashboards for use by business users and executives

The ultimate goal of a MDS is to accelerate the analysis of a company's business data and then quickly generate information employees can use to create process efficiencies or uncover opportunities. It does that by operationalizing the movement of data in both volume and direction into a central repository (the data warehouse/lake) and from there into an analysis platform.

By doing so, the modern data stack greatly increases the speed with which employees can generate insights out of data — without requiring companies to rip and replace their existing infrastructure.

In other words, reports that used to take weeks or even months to trickle down to frontline workers can now be generated and shared in hours or minutes.

The cloud can take much of the credit for simplifying and accelerating the process. Because the MDS is hosted in the cloud, it's much easier to scale quickly and is accessible to people regardless of location or their device.

Cloud services also provide a plethora of proven and rich algorithms and technologies that companies can tap into whenever they need help solving data issues. They can do this without having to worry about keeping these technologies current or maintained — the cloud provider takes care of that.

Compare that to the days when the only choice was on-premises servers, when organizations were required to buy off-the-shelf software and then figure out how to make it work for them.

For a lot of data shops, the ability of the cloud and MDS to provide at-the-ready tools and solutions and accelerate the entire problem-solving process is very attractive.

Modernizing your data stack, step by step

The first thing to know about implementing a modern data stack is that everything is hosted in the cloud and all the labor and headaches involved with building infrastructure don't apply.

Here are the steps involved:

- First, decide which data warehouse you're going to use to consolidate your data. The top contenders among data warehouses include Snowflake, Amazon Redshift and Google BigQuery. A warehouse must provide the easy scaling of computing and storage duties while avoiding extended downtime. A comparison of some of the players can be found **here**.
- Second, select a data pipeline technology to move data into the warehouse. Ideally, you will need a fully managed ELT data pipeline that eliminates the need to spend time worrying about API and schema changes. A well-equipped ELT solution also comes with pre-built connectors to each of your company's important data sources, and ELT must enable the rapid scaling of data integration.

- Choose a transformation technology that works well with the chosen data warehouse and supply features that simplify the tracing of data lineage, such as version control and documentation that assists in outlining the impact of transformation on tables. It's helpful when it's written in a language analysts and engineers understand, such as SQL.

- The machine learning algorithms offered by the leading cloud providers are one of the most attractive features of the cloud and one of the reasons so many companies have adopted MDS. In addition to vast machine learning libraries, your provider may also offer pre-built predictive models for certain scenarios. These features save companies enormous time and costs compared to the traditional way of cleaning data, building models, testing and then applying them. Some of that is still necessary, but the cloud makes this a lot less time consuming. And again, these cloud services run on optimized (scalable) environments — often GPU powered.

- Finally, select an analytics platform for your end users to use. An abundance of self-service analytics technologies are also available in the cloud. An example of this is Looker (acquired by Google in 2019), which produces tools for data exploration and business intelligence. Late last year, **Google also partnered with Tableau**, so Google Cloud Platform can now access Tableau's dashboards and other data visualization tools.

Most companies have already begun migrating at least some of their data to the cloud. But how do you decide when it's right to upgrade to the MDS?

Each company must consider the value of such a transition as well as the risk of waiting and possibly falling behind competitors.

When all the benefits offered by an MDS are factored in — the elimination of technical barriers, the scalability and end-user accessibility, and the ability to leverage readily available machine learning algorithms — it's difficult to conceive of a scenario where a company doesn't win in the long run.

Putting it all in motion

It's not realistic to expect CDOs to create a complete overhaul of a company's data infrastructure in less than two years.

What they can do, however, is streamline the operational side of the company's data stack and post some early, significant wins. Companies determined to see rapid change can adopt technologies that help them manage data more efficiently and cost effectively while yielding better results.

CDOs determined to effect rapid change will make the best use of the cloud and data management technologies, which can eliminate much of the complexity and provide the kind of momentum needed to score some early victories. Even better, the right technologies will also lay the foundation for subsequent achievements, allowing that momentum to continue.

Data is too important to the process of decision-making at modern companies to rush the creation of a well-equipped and stable analytics strategy. To that end, companies should change their thinking about costs, and think more in terms of investments. As noted earlier, the costs associated with ushering in a data-driven business transformation is often the reason some companies scuttle their initiatives.

Data leaders should look at the cloud and the modern data stack for what they are:

- Opportunities to increase the value of data and turn it into a powerful competitive advantage.
- The money spent on data-powered transformation is a smart investment.
- CEOs and boards should also keep in mind that an MDS is not just a set of tools but a process.

A few tips for CDOs: Don't be shy about promoting the triumphs of your team and spend time helping your company's leaders understand the return on their investment. Your counter parts in marketing have long illustrated their contributions to the company by providing illuminating metrics.

KPIs should tie to quantifiable business milestones rather than technical accomplishments to ensure all data initiatives are business-focused. Common categories include:

- Operations: Enabling data usability, availability and efficiency
- Analytics: Supporting analytics and reporting on products, customers and business operations
- Governance: Monitoring and governing enterprise data
- Innovation: Monetization of data and managing costs and SLAs

Next, focus on constructing a modern data stack. The MDS hands you the means to leverage all your existing data infrastructure, score early wins and lay the groundwork for numerous transformational data-driven moves to come.

Finally, don't do it all yourself. Build relationships, and reach out to people and services with the expertise and the technologies that help you keep costs down while helping to ratchet up your ability to generate insights for your company.

If you're a data leader, reducing time to value should be top of mind. You need to demonstrate the transformative power of data — but it's probably safer to pick your battles carefully at least early on.

This bears repeating: begin by finding an issue that has frustrated the company but can be improved or eliminated by data analysis.

Then identify the data sources and find the right technology partner, one that can help unlock the relevant data from your sources (and others you anticipate needing to unlock data in the future). Accomplish that and you should be on your way.

Fivetran is the global leader in modern data integration. Our mission is to make access to data as simple and reliable as electricity. Built for the cloud, Fivetran enables data teams to effortlessly centralize and transform data from hundreds of SaaS and on-prem data sources into high-performance cloud destinations. Fast-moving startups to the world's largest companies use Fivetran to accelerate modern analytics and operational efficiency, fueling data-driven business growth. For more info, visit **Fivetran.com**.