



The New Leadership Mindset for Data & Analytics

CONTENTS

EXECUTIVE GUIDE

The New Leadership Mindset for Data & Analytics

1 How Large Companies Can Grow Their Data and Analytics Talent

By Thomas H. Davenport

4 Changing Culture Is Central to Changing Business Models

By Lanham Napier, Barry Libert, and K.D. de Vries

7 Why Chief Data Officers Must Assume Leadership for Data Success

By Randy Bean

11 Top-Down Leadership for Data: Seven Ways to Get Started

By Thomas C. Redman

14 Leading With Decision-Driven Data Analytics

By Bart de Langhe and Stefano Puntoni

18 Empowering a Data Culture From the Inside Out

Jonathan Tudor, interviewed by Ally MacDonald

21 Developing the Data-Driven Organization: Leadership, Culture, and Learning

Sponsor's Viewpoint from AWS



How Large Companies Can Grow Their Data and Analytics Talent

Companies need to identify the type of talent they need in order to become data-driven.

BY THOMAS H. DAVENPORT

While many companies are hiring data scientists and other types of analytical and artificial intelligence talent, there is little consensus within and across companies about the qualifications for such roles. The term *data scientist* might mean a job with a heavy emphasis on statistics, open-source coding, or working with executives to solve business problems with

data and analysis. The idea of data scientist “unicorns” who possess all these skills at high levels was never very realistic.

As the job has grown more popular and sought-after, an increasing number of professionals have begun to use it to describe their role. Colleges and universities have responded to the demand as well by offering hundreds of new programs on data science and analytics. But the skills taught in such

programs vary widely, and some universities offer multiple programs with different emphases. For both newly hired and experienced employees, titles such as data scientist and quantitative analyst are not likely to be good guides to their actual capabilities.

While there are initiatives to standardize the different types of data and analytics roles and requisite skills across organizations, they are in the early stages. The idea

behind these initiatives is excellent, but developing new standards typically takes many years.

In the meantime, companies need to devote considerable attention to classifying and certifying the different types of analytical jobs they have and need. They also would benefit from expanding their talent pools by working with universities directly on educational programs and by building and nurturing communities within their organizations to develop employees for their data teams. These steps are essential for companies looking to use analytics to improve both operations and opportunities for digital innovation — companies such as TD Bank Group.

Understand the Data Roles Needed

Analytics and AI talent is a scarce and valuable resource for every company, but particularly for those with a desire to be data-driven. That's the situation at TD Bank Group, a large North American bank (the largest by assets in Canada). TD made the decision to use its vast amounts of data to power new data-driven customer experiences and improve its own processes. To do so, it has embraced data, analytics, AI — and the need to nurture the talent that will use those tools throughout the organization.

TD executives made a series of internal investments directed at the critical importance of assessing, hiring, and motivating its data and analytics talent. Peter Husar, vice president of enterprise analytics strategy and planning, told me that he and his colleagues saw an opportunity to expand the role of analytics by first bringing consistency to the job definitions and positions related to data and analytics that existed across the bank.

Husar and his team focused on understanding who was in their data and analytics community in order to provide a clear picture of their talent landscape. They worked quickly to create standardized job groupings and classifications by identifying seven job families (roles that focused on advanced analytics, business application management, information management, insights, business intelligence and reporting, data governance, and visualization).

Within each job family, specific jobs were defined in detail — in all, 65 different roles across seven job families. For each role, several attributes were

described, including its primary purpose, its key accountabilities, its breadth and depth, and the experience and education required to perform it. Individual contributor attributes were identified, as well as attributes for people managers. Additional competencies were labeled across the job classifications, along with competency assessments and a self-assessment process. With this new framework, TD employees working in data and analytics roles were then mapped to the job families and specific jobs. TD has an Office of the Chief Data Officer, which oversees enterprise data management, analytics, and AI.

Although most companies don't have the scale and scope of analytical talent that TD does, many medium to large companies have some data scientists and quantitative analysts. In order for their projects to succeed, it's important to create at least a simple

classification of their capabilities. Are they primarily oriented to statistical analysis and modeling, coding IT solutions, or translating business issues into data and analytics solutions?

Projects may require all three skills, but those skills are unlikely to be found in one person.

Acquire and Retain New Talent

In addition to more effectively managing existing data and analytics talent, organizations need to make a concerted effort to bring on and retain highly skilled employees.

In 2018, TD acquired artificial intelligence company Layer 6 and made one of its cofounders, Tomi Poutanen, the chief AI officer for the bank. TD continues to transfer Layer 6's capabilities to the bank at large — including information-sharing programs at Layer 6 for data and analytics employees.

The bank has built strong relationships with a number of Canadian universities, including the University of Toronto's Rotman School of Management, Western University, Queen's University, and York University's Schulich School of Business. TD sponsors events at these institutions, including hackathons and speaker series. At Rotman, TD has funded both the TD Management Data and Analytics Lab and the Rotman Financial Innovation Hub in Advanced Analytics to help foster the development of data and analytics professionals.

The idea of data scientist 'unicorns' who possess all these skills at high levels was never very realistic.

TD also works closely with leading AI research programs in Canada, including those at the Vector Institute in Toronto, of which TD was a founding sponsor, and Mila in Montreal, in order to foster partnerships and connections.

These efforts have not gone unnoticed. TD has been rated among the most attractive employers for data and analytics talent by Canadian students: In a large student survey in Canada, analytics students ranked TD as one of the top four companies they would consider working for.

Most organizations will have to turn to multiple suppliers of talent to acquire the range of skills needed. They should work with the universities and colleges that are providing them with new talent to better understand the foci of the schools' data science and analytics programs. They should also embark on using job standardization and classification to design internal education and skills development programs. When organizations make clear which skills are associated with which jobs, employees can more easily decide which roles they want and acquire the skills to perform them.

Build a Data and Analytics Community

TD data and analytics leaders are conscious of the need to build community and knowledge-sharing approaches among the data and analytics professionals across the bank. That includes providing opportunities for data- and analytics-oriented employees to network and hear about trends in the industry. Five years ago, the bank had its first annual Big Data & Analytics Summit, with 40 employees attending. In 2019, the bank had its fifth annual summit (disclosure: I was a paid speaker at this event), with over 2,000 employees in attendance — and the numbers continue to grow.

Another key aspect of community is volunteering. An initiative called TD Mindpower: Analytics for Social Good was created in 2019 to pair data and analytics employees with data and analytics projects for nonprofits in Canada and the U.S. As part of the bank's acquisition of Layer 6, a pledge was also made to dedicate time and talent to work specifically on AI projects for social good.

TD's initial job-mapping exercise led to a clearer vision for how individuals can move around, upskill, and grow their career paths. The team spearheading the plan created educational materials to help

data and analytics employees develop their skills and made those materials available on an enterprisewide online platform for self-learning and development. The bank is now focused on expanding an understanding of the potential of data and analytics to all TD employees and has developed learning curricula for nonanalytics professionals as well. The goal is to further embed data-driven decision-making in all aspects of the bank.

According to Husar, data, analytics, and AI at TD have come a long way in the past five years and are helping to shape and deliver tangible results for customers and employees. In particular, during these unprecedented times, the bank has shown its ability to pivot, leverage data and analysis, and mobilize quickly to respond to the needs of its customers, spinning up new and innovative ways to stay connected and answer questions. For example, TD has used customer data and analytics to personalize customer communication. Customers who visit often received targeted messages about changes in bank branch hours, and those most likely to have received a stimulus check got messages about how to make mobile deposits.

Whether people are called data scientists, quantitative analysts, or something else, these titles are still relatively new to companies. It's not surprising that there is little consensus on their meaning. Over time, new titles will emerge that better describe the particular activities involved in a job; data engineer, for example, has already arisen to describe people who spend most of their time wrangling data. Until more consensus is reached across our society, however, it's important for companies to emulate TD Bank and clearly identify the types of talent required to become data-driven and then acquire it, nurture it, and unleash it.

Thomas H. Davenport (@tdav) is the President's Distinguished Professor of Information Technology and Management at Babson College, as well as a fellow at the MIT Initiative on the Digital Economy and senior adviser to Deloitte's Analytics and AI practice.

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.

Changing Culture Is Central to Changing Business Models

Leaders need to examine their core beliefs if they want to prosper in a COVID-19 world.

BY LANHAM NAPIER, BARRY LIBERT, AND K.D. DE VRIES

Culture is all-encompassing. It radiates through every action taken inside an organization — including deciding what is made and sold, which employees are hired and retained, which customers are serviced and how, what is measured and reported, and where time and money are invested.

However, while some leaders strive to create tomorrow's cultural norms using advanced technologies, others resist cultural changes and stick to their knitting. The result is a growing chasm between winners and losers. For example, today's trillion-dollar behemoths are all technology-based organizations that take advantage of mobile technology, data, multiple revenue streams, and AI strategies. They are led by leaders who want to change the world, while the companies that held the top spots of growth and value just two decades ago — banks, oil companies, real estate companies, and manufacturers — are paying the price for holding on to their historical cultural beliefs.

A good example of what's at stake is the emergence of electric vehicles and Tesla's rise to prominence as the most valuable car company in the world. In the companies that Tesla has surpassed, boards and leaders have resisted adapting to new cultural and technological realities. It's not that the emerging cultural norms about the advantages of embracing modern business models have been unclear. It's that while culture provides the foundation for organizational and industry stability, it is also the force that keeps leaders stuck in their old ways of conducting business. The takeaways from this story apply to every industry.

Leaders who are serious about creating the organizations of tomorrow have a simple choice: They can stay with the cultural norms that created their prior success, or they can do the hard work to change themselves to ensure success in the future. Today's leaders need to take a personal journey to avoid the fate that has befallen companies such as Blockbuster, Kodak, Sears, and so many



others, and it starts with three steps: examining personal values in order to redefine them, communicating the new values widely, and measuring what matters — the performance of the new initiatives and investments that are necessary.

Leaders: Change Yourself to Change Your Culture

Cultural norms are deeply held beliefs about the way an organization should work. These norms translate into different cultural products throughout the company, including values, customs, and traditions. Said differently, cultural norms define an organization as well as its sources of growth and value.

For example:

- If leaders believe in manufacturing, they make things and measure units produced.
- If leaders believe in services, they deliver support and measure hours billed.
- If leaders believe in analytics and AI, they measure data collected and insights generated.

But new technologies — coupled with consumers' changing wants and needs — disrupt legacy beliefs. Twenty years ago, most leaders would have said they believed that making things was more important than matching buyers and sellers. That core belief has been shaken to its core, of course, as subscription and marketplace models powered by data and AI began to fuel the most powerful companies in the world today — Amazon (with Prime and its massive supplier network), Apple (with its iOS developer community), Facebook (with its billions of users), and Google (with its search and matching algorithms). Further, external forces like the COVID-19 pandemic can accelerate the demise of traditionally held norms and make those who are best prepared for the future the winners of today (such as Zoom, Shopify, and Spotify).

While culture provides the foundation for organizational and industry stability, it is also the force that keeps leaders stuck in their old ways of conducting business.

Three Steps Toward Change

If leaders are itching to change their corporate culture and direction, we recommend that they first follow three steps to challenge their own biases:

Step 1: Leaders need to examine their thoughts and values in order to redefine them. Attempts to change an organization's strategy, products, services, measurements, or reporting will be in vain if leaders don't realize that cultural change is an inside job. Leaders have to change first.

Take Ford, for example. Along with other car manufacturers, Ford has finally decided to enter the electric car race. However, for Ford to catch Tesla, its leaders need to let go of the company's rigid traditional management practices that made it successful. That's the only way Ford will be able to compete for customers as other U.S. states join California in prohibiting the sale of internal

combustion vehicles and, in so doing, create the new cultural norms for the industry. This is in stark contrast to the days when Henry Ford started his business and became the cultural trendsetter.

Before they look at their products, people, and processes, leaders have to better understand their own underlying attitudes, behaviors, and beliefs. Only by challenging and changing core beliefs can leaders consistently act in ways that support a true business model transformation.

Step 2: Leaders need to communicate their new identity. Clearly defining the values that management and the board hold should help an organization create today's newest business model: AI-powered digital platforms with multisided revenue models. Doing this requires that they discourage existing employees from keeping their heads down and proceeding with business as usual.

A stunning example is today's colleges. For a long time, universities' leaders resisted restructuring their business model — by sticking to a primarily on-premises structure despite its cost to students, limits on access, and restricted extensibility. COVID-19 is forcing a sea change: This fall, at least 65% of higher-learning institutions are holding classes completely or partly online, according to the College Crisis

Initiative at Davidson College. The astonishingly fast shift to digital classes will drive down the per-student cost of traditional instruction if it sticks as the pandemic wanes. Colleges and universities will need to rethink their identities and the new value they are proposing to justify their high price tags. This work should include examining how data and AI can be leveraged to reinvent the college experience.

Company employees, customers, investors, and lenders all need to understand a leader's new belief system before they will be ready and willing to align their actions. Only by overcommunicating can leaders ensure that their new belief system is accepted and embraced.

Step 3: Leaders need to measure the impact of new cultural norms on the company's performance. To truly transform, leaders must identify new key performance indicators that link to

their company's emerging new identity and track and report them. If you don't measure it, you won't manage it, nor will you value it. Leaders and boards that continue measuring the same things they have always measured and valued will end up with what they already have — a slow-moving legacy business. But if they do want change to stick, they will need to follow up with adjusted compensation and rewards.

For example, Blockbuster measured brick-and-mortar stores and individual video rentals, while Netflix measured and built subscriptions (first with a physical product and then, explosively, through a digital platform). John Antioco, former CEO of Blockbuster, recognized the importance of heavily investing in a digital platform. But at the first hint of unprofitability during this investment in the transformation of his company, the board and investors quickly reverted to their old beliefs about what creates value and where to allocate dollars, and Antioco was ousted. Within a few years, Blockbuster declared bankruptcy.

Leaders often fail to change what they measure, manage, and report, hoping that somehow those parts of the system will be OK while the rest of the organization changes. But the best thing leaders can do to change behaviors is to change what is measured, managed, and reported.

Change Your Leadership Culture — or Your Leaders

Organizational culture has a strong impact on the efforts of an organization trying to adopt big data, machine learning, and network-based business models to catch up with today's leading companies. The reinforcing loop — culture — is the foundation of that race. If leaders truly want to derive meaningful business benefits from analytics and platform models, they must proactively address their own core identities before trying to introduce large-scale transformation initiatives.

Most leaders simply don't want to put in the work and examine their core beliefs, and neither do their boards. The fear is that they will become destabilized. As a result, both leaders and boards make short-term decisions to remain aligned with legacy industry norms that in the long run lead to their demise. Consequently, sometimes the only way out is to replace leaders, including the board (and, potentially, investors, too).

To paraphrase famed U.S. baseball manager Yogi Berra: You've got to be careful if you don't know where you're going, let alone what you stand for. Otherwise you might not get there.

Lanham Napier (@lnapier) was the CEO of Rackspace and is the CEO and cofounder of BuildGroup, a growth equity fund that invests in public and private SaaS-plus marketplace companies. **Barry Libert** (@barrylibert) is the CEO of AIMatters, a business model data science startup, and a strategic adviser to BuildGroup. **K.D. de Vries** is a marketing consultant at BuildGroup.

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.



Why Chief Data Officers Must Assume Leadership for Data Success

Chief data officers need to be equipped with the right tools and support so they can execute on the company's data vision.

BY RANDY BEAN

In their efforts to become data-driven, Fortune 1000 companies face the common challenge of identifying the right leadership that will enable them to overcome cultural and business roadblocks. For most businesses, cultural issues manifest themselves in a variety of ways — resistance to change, antiquated business processes, a lack of

clear coordination and communication of business imperatives and business value, ineffective organizational alignment, and uncertain leadership and commitment to data initiatives.

These issues all point to serious gaps between ambition and execution that most organizations confront when embarking on data transformation efforts. As companies

struggle to manage data as a vital business asset, they must develop the leadership skills, expertise, and organizational structure to effectively manage and communicate the business value of data. Now more than ever, it is incumbent upon them to establish strong data leadership that will define and deliver on a data vision that supports the greater business vision of the company.

The Emergence of the Chief Data Officer

In the wake of the 2008-09 financial crisis, leading financial services companies were compelled by government regulators to address systemic issues resulting from failures in the management and reporting of financial data that were seen as a contributing factor to the financial collapse. In response, several large banks took the lead in creating the role of chief data officer (CDO) to focus on implementing standards and safeguards for managing data. The driving motivation was to mitigate risk and avoid future systemic financial failures.

The rise of the CDO as a corporate risk function happened in step with the emergence of big data in the early 2010s. Although a few intrepid organizations had designated executives to function as CDOs earlier, a new combination of *defensive* drivers (compliance, risk mitigation) and *offensive* drivers (revenue generation, business growth), based on the promise of big data, resulted in a sudden and dramatic increase in chief data officer appointments. The percentage of major companies with a CDO rose from 12% in 2012 to a peak of 67.9% by 2019, according to an annual executive survey by NewVantage Partners.

Today, the chief data officer role has emerged as a standard for most Fortune 1000 companies, but it comes with serious issues and challenges as companies struggle with how best to shape the role to achieve successful business outcomes. While a majority of CDOs — 54.6% — are now focused on revenue initiatives (offensive), a significant minority — 45.4% — remain focused on risk factors (defensive). Only 12.3% of CDOs have assumed direct revenue responsibility thus far, suggesting that moving into an offensive business-generation role will take some time.

Alarming, less than 30% of executives report that the CDO role is successful and well established within their organizations. It should therefore not be a surprise that the CDO role has one of the highest turnover rates in the C-suite, especially when compared with more established positions like the CFO or CIO. Several major banks are on their fourth or fifth incumbent in the CDO role, and others have

paused to rethink the role, highlighting the challenge that CDOs face.

The Need for Business-Driven Data Leadership

Data challenges remain an issue for most companies — with only 37.8% reporting that they have established a data-driven organization, and only 26.8% saying that they have forged a data culture. Because there has not been an established tradition or consensus on how data should be managed as a business asset, most organizations are charting new territory as they go. Companies are actively seeking leadership to move their organizations forward in an area where they continue to struggle and where most companies remain at an early stage of maturity.

Among the constituencies that have been grappling with and thinking about data leadership for the longest time is the MIT Chief Data Officer and

Information Quality (CDOIQ) Program, led by Richard Wang, who previously held the role of deputy CDO and chief data quality officer for the U.S. Army. He was instrumental in spearheading the launch of the MIT

CDOIQ Symposium, which is now in its 15th year.

I have had the good fortune to host and moderate a panel focused on the CDO role at the symposium for the past six years. In 2020, due to COVID-19, the event was held virtually for over 1,000 participants, approximately 25% of whom identified themselves as current CDOs. This year I was joined by CDOs from American Express, Bank of China, Bristol-Myers Squibb, CVS Health, and Mastercard, who shared their perspectives on data leadership and their recommendations on specific steps organizations can take to ensure data success:

1. Develop and execute a data strategy that mirrors your business strategy and vision.

Mastercard CDO JoAnn Stonier advised that for companies to be data-driven, “the data strategy of a company must follow the business strategy.” Since each organization is different, embodying its own business objectives and culture, companies must commit to a data strategy that reflects the relevant use cases, supply chain of data, and data consump-

Organizations that take steps to build a robust data organization from top to bottom will have an advantage in establishing a data culture.

tion patterns that drive business outcomes particular to them. As American Express CDO Danielle Crop put it, “Data is like air. It is everywhere.” Value is derived from data only after it is organized, analyzed, synthesized, and presented. Data becomes relevant in the context of the business strategy and goals. Bank of China CDO Richard Goldberg concluded that to be effective, CDOs must support the business.

2. Relentlessly communicate the context, complexity, and value of data up and down the organization. Data brings value to the extent that its benefits are clearly articulated and understood throughout the organization. Many data initiatives fail to succeed due to the complexity of managing data and the failure to communicate business value in clear ways that cut across business and technical boundaries. For data leaders, communication is key: Speak in business terms, and avoid technical jargon.

In our panel, Mastercard’s Stonier explained, “Data is an intangible representation of the world around us. How you use data and the insight that comes from data is the exciting part.” A CDO must be multilingual — that is, able to speak the language of the technologist and the data scientist and be able to translate it into the most fundamental business terms and business value for others in the organization. “There needs to be constant education to demonstrate the value of your work, reinforce the message, and communicate the imperative of data,” Goldberg noted.

3. Boldly engender trust in your data to assure business and customer buy-in, credibility, and momentum. Many data initiatives fail because business constituents lose faith in the data. Factors that might lead to such distrust include incompleteness, inaccurate or misunderstood representation, or a lack of common understanding and definitions. American Express’s Crop observed, “The future of the CDO role is about enabling trust in data and driving innovation. CDOs must engender that trust in data.”

Consumers, executives, and the public must be able to rely on the integrity, quality, accuracy, and transparency of data to make decisions. Former Bristol-Meyers Squibb CDO Krishna Cheriath of-

fered a way leaders can frame this by asking, “How do we create a culture of digital citizenship to ensure that we are using data responsibly?”

“Data discipline must be part of the enterprise agenda,” Cheriath added. It is the mandate of the CDO to build and sustain this trust in data across the enterprise and among the constituencies of the business.

4. Create a compelling career path for data leadership within your organization. Bank of China’s Goldberg described the “need to create a career path to enable the business to embrace data.”

This entails building systemic capabilities among all functions relating to data, including strategy, governance, architecture, analysis, and consumption. Goldberg noted that

by executing consistently, an organization will come to “embrace data as a business asset.” Bob Darin, CDO of CVS Health, observed that coordinating across the enterprise is one of the biggest challenges for the CDO and noted that his role “entails a significant amount of change management — it is the biggest part of the job.”

Organizations that take steps to build a robust data organization from top to bottom will have an advantage in establishing a data culture.

The Data-Driven Business Imperative

Companies cannot be successful and survive over time if they are not effective at managing risk, serving their customers, and identifying new ways to grow and evolve their businesses over time. Those companies that have learned to harness data, particularly as part of the digital experience, have been rewarded with rapid growth, customer expansion, and dominant market share and market value. Success is not assured. The long-term outlook for the role of the chief data officer remains uncertain. In 2019, 17.5% of survey respondents suggested that the CDO role was unnecessary and would ultimately be phased out over time as organizations look to data leadership from other C-level functions or as the need diminishes as companies became more data-driven and self-sufficient. It has been noted that data-driven companies like Amazon and Google have not traditionally appointed CDOs because data is already deeply ingrained in their DNA.

For data leaders,
communication is key:
Speak in business terms,
and avoid technical jargon.

The establishment of the chief data officer role represents a recognition that data matters to the organization. And organizing, governing, and delivering value from data can be one of the primary factors that differentiate those companies that seize the moment from those companies that lose their way, especially in the digital economy of the present and future. Companies must ensure that the chief data officer has the necessary tools and support for executing on his or her data vision. Only then will businesses be able to legitimately claim that they have earned the right to be called *data-driven*.

Randy Bean (@randybeanvp) is an industry thought leader, author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm that he founded in 2001. He is a contributor to MIT Sloan Management Review, Forbes, Harvard Business Review, and The Wall Street Journal. Contact him at rbean@newvantage.com.

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.

Top-Down Leadership for Data: Seven Ways to Get Started

Leaders can initiate successful data strategies by focusing on data quality, building organizational capabilities, and putting data to work in new ways.

BY THOMAS C. REDMAN

For many senior executives, data presents a conundrum. Naturally, they want their data programs to succeed. Indeed, they'd like to help in some way, and even provide leadership, but beyond providing funding, they're not sure how. Business leaders' lack of technical knowledge exacerbates their uncertainty and makes them feel like outsiders to data science teams.

But data programs cry out for business leadership, and there are many ways senior executives can play a bigger role, even if they don't fully understand data. This article explores seven ways leaders can accelerate their companies' data efforts, derive near-term benefits, and gain a better understanding of the roles data can play in advancing their business objectives. Think of these as opportunities. Not all of them will suit your needs, interests, or style, so just pick one or two and get started.

Focus on Quality

We begin with three opportunities related to quality. Counterintuitively perhaps, improving data quality saves money. Further, all data strategies (and more and more business strategies) depend on high-quality data.

1. Break the quality logjam. Many leaders readily admit that they don't trust their company's data, at least not enough to use it when it really matters. It's no wonder, given how often they encounter bad data: Two numbers don't jibe, some report looks too good to be true, and the finance team complains every quarter about the overtime needed to produce a routine report. Still, most companies manage to convince themselves that their data is "pretty good." And an uneasy stasis results — you don't have the trusted data you need but are frozen from doing anything significant about it.

To break the logjam, undertake a review to determine how good — or bad — your data



is. Tell people there will be no rewards or punishments for good or bad results — it is just important to establish a baseline. Instruct your teams to make a simple data quality measurement, using what I call the Friday Afternoon Measurement. It entails laying out a sample of the most important data your teams use every day, marking the obvious errors, and then counting them up. This measurement returns a score between 0 and 100 — the higher the better.

Although calculating this measurement takes just a few hours, it may take longer for the team to work through the implications. On the one hand, you may find out that your data can be trusted (scoring above 96), and you can then move on to other improvements. However, few teams score this well, and any score below 85 signals real problems. Solving these problems will require more work, but breaking the logjam is half the battle.

2. Get to the bottom of something. When you raise questions about numbers that just don't look right, your staff goes off in a frenzy to find the mistakes and correct them. Tragedy averted, or at least delayed.

This little drama plays out dozens, hundreds, or thousands of times a day (depending on the size of your company) — sales cleans up data it receives from marketing, operations cleans up data from sales, and finance cleans up data from everyone. Note the vicious cycle and the incredible drain on company resources. You can lead by example to break this cycle. Next time an issue comes up, lead an improvement project to get to the bottom of it — finding, then eliminating, the root causes.

In so doing, you will send a signal that your company must change its approach to data quality, from passively dealing with errors to making them go away — for good. Those who report to you (and maybe some peers) will follow suit, making this perhaps the single most transformative action you can take.

3. Get everyone on the same page. People in all but (maybe) the smallest companies complain, “Our systems don’t talk.” You’ve probably experienced this yourself, when people cite it as the reason why different answers might exist throughout the company to basic questions like “How many cus-

tomers do we have?” Unfortunately, the problem of systems not talking camouflages a more fundamental issue — namely, people not speaking the same language. This has real consequences: It is more difficult to work across departments, complexity grows, and your technology department wastes time better spent elsewhere.

To find out if you have this problem, pick three or four important terms — for a financial services company, *security*, *buyer*, and *client* are good examples. Ask everyone (yourself included) to write down their definitions of these terms and bring them to your next staff meeting.

Then read them aloud and see how closely the definitions align. Remember, there is no right or wrong. If everyone agrees, great! If not, you may have a serious problem.

This is the headiest of the seven opportunities described here. But once you know where to look, you will see how common language can resolve all sorts of business issues.

Put Your Data to Work

4. Bring data science to your strategic decisions.

A common misconception about data science — with terms such as *big data*, *artificial intelligence*, and *machine learning* — is that it's reserved for areas where data is plentiful. In reality, when done well, data science helps ensure that a business problem is clearly articulated, reveals hidden biases in your thinking, provides a fresh perspective on whether the data you have is good enough to address that problem, seeks new sources of data to close gaps, analyzes all the data in powerful ways, and reduces your uncertainty. So, pick an issue, invite a first-rate data scientist to join the fun, treat the person as a full member of your team, and listen to what he or she has to say.

5. Determine how your data sets you apart.

Your next opportunity involves competitive differentiation. As you know, companies don't compete based on the ways in which they are similar to others but on the ways in which they differ from the crowd. It is all too easy to view your data as a byproduct of your real work and not see it as a source of value in its own right. It is worth the time to examine where strategic opportunities may be overlooked.

Send a signal that your company must change its approach to data quality, from passively dealing with errors to making them go away — for good.

Start by charging a multidisciplinary team with answering the question, “Do we have data that no one else does?” Chances are, you do — after all, your data is uniquely your own. If so, this data qualifies as proprietary and may be a source of competitive advantage. Clearly enough, the next step involves figuring out how to exploit it. Of the opportunities described here, proprietary data probably signals your best chance at creating near-term revenue.

Build Needed Organizational Capabilities

6. Separate the management of data and technology. Slowly, and usually with great difficulty, data is invading every nook and cranny of every industry, company, and department, creating both opportunity and risk. Yet today’s organizations are unfit for data — they lack talent, silos get in the way of data sharing, data and business strategies are poorly connected, and it is unclear who is responsible for what in the data space. Over the long haul, these issues will require a lot of your time.

Far and away the best thing senior leaders can do for their data programs is to clarify managerial accountabilities. Fortunately, step one is relatively straightforward: Move lead responsibility for data out of your IT department. Data and technology are different kinds of assets, and comingling their management has slowed progress on both. Work department by department, from HR, to finance, to operations, to administration, to find a better home for your data program.

7. Demonstrate your commitment to data on your board. It’s important to mirror the data improvements you make internally on your board of directors. If you haven’t done so already, make it a priority to appoint a director to your board who understands data. You need someone who will push you to go further and faster than you might otherwise be inclined to, to serve as a sounding board as you evaluate options, and to help you anticipate and deal with resistance. The ideal candidate has both an expansive vision and experience with hard-fought battles in advancing a data agenda. The sooner you find this person the better, so start looking right away. Interview as many people as you can, and don’t decide until you find someone you really trust.

Data may well represent your best chance to grow your business and distance yourself from your competitors, but getting even a fraction of the value

data has to offer is tough. It’s no surprise that unlocking this potential requires strong stewardship from senior leadership. Executives can use the seven ideas in this article as entry points for getting started.

Thomas C. Redman (@thedatac1) is president of consultancy *Data Quality Solutions* and author of *Getting in Front on Data: Who Does What* (Technics Publications, 2016).

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.

Leading With Decision-Driven Data Analytics

Data analysts often fail to produce insights for making effective business decisions, but that's not their fault. Leaders need to make sure that data analytics is decision-driven.

BY BART DE LANGHE AND STEFANO PUNTONI

If you were to ask any major CEO about good management practices today, data-driven decision-making would invariably come up. Companies have more data than ever, but many executives say their data analytics initiatives do not provide actionable insights and produce disappointing results overall.¹

In practice, making decisions with data often comes down to finding a purpose for the data at hand. Companies look for ways to extract value from available data, but that doesn't necessarily mean data analysts are answering the right questions. It's also not a safeguard against the influence of preexisting beliefs and incentives.

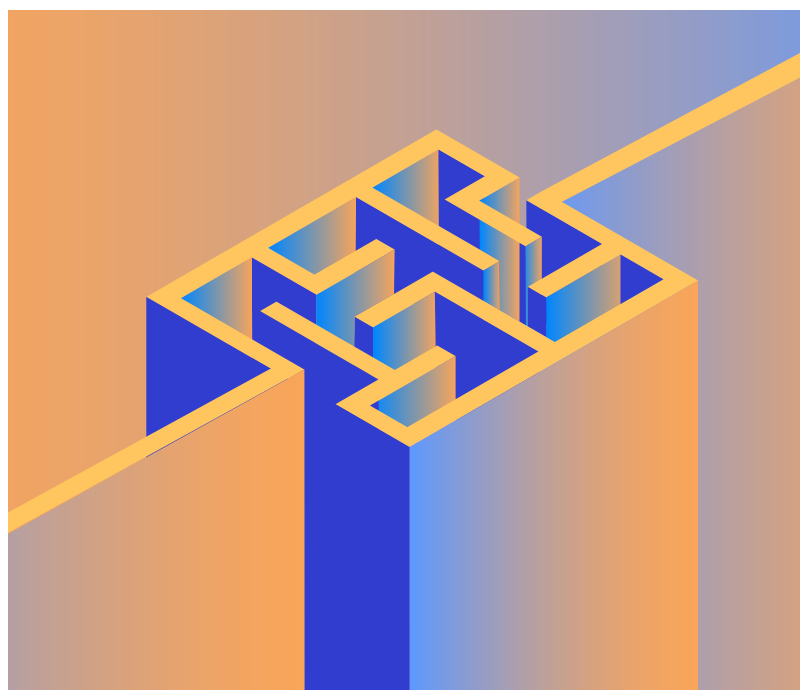
The solution is simple: Instead of finding a purpose for data, find data for a purpose. We call this approach *decision-driven data analytics*.

'Data-Driven' Often Means Answering the Wrong Question

We'll use some example companies to make our point. Let's start with customer relationship management at RollingBoulder, a media company with a subscription-based business model. RollingBoulder's customers can renew their annual memberships by responding to a renewal letter that they receive when their membership is about to expire. To reduce customer attrition, the organization sometimes adds a thank-you gift to these mailings.²

Over the years, RollingBoulder has developed a rich data set that describes past and current customers along various dimensions, such as location, membership duration, and website usage and behavior. The company has developed a sophisticated predictive algorithm that uses this information to quantify the likelihood that an active member will churn, and it then sends gifts to at-risk customers.

This data-driven approach to churn management is considered a best practice in the industry, but it is flawed.³ To understand why, let's look at the central question the company is addressing with this analysis: "How likely is a customer to churn?" This is valuable information. For example, it allows the



company to make projections about the value of its customer base. However, it does not address the question that is relevant here: “What is the effect of including a gift on a customer’s likelihood to churn?” This question cannot be answered based on data the company has already gathered and requires further data collection and analysis.

Data-driven decision-making anchors on available data. This often leads decision makers to focus on the wrong question. Decision-driven data analytics starts from a proper definition of the decision that needs to be made and the data that is needed to make that decision.

‘Data-Driven’ Often Means Reinforcing Preexisting Beliefs

Take the case of Gwenn & Jenny’s, an ice cream vendor that wants to know how its online advertising on Twitter impacts sales. The social media platform offers a three-step attribution process to evaluate the sales impact of advertising.⁴ First, a data broker, like Datalogix, shares identifying information from Gwenn & Jenny’s customers (like browser cookies, email addresses, and phone numbers) with Twitter. Next, Twitter searches for those customers in its records and, if there is a match, adds information about these customers’ activities on the platform (like whether they viewed or clicked on the brand’s tweets). Finally, analysts compare purchase decisions of customers who engaged with the Gwenn & Jenny’s brand on Twitter with purchase decisions of customers who did not.

This approach invariably reveals stark differences for a brand: Customers who saw and engaged

with a brand on Twitter visit its stores more often, and they spend more on each visit. Taking the view that this data suggests social media advertising has a large impact on sales supports Twitter’s business model. It’s also consistent with clients’ beliefs that social media advertising works and that its effectiveness can easily be measured.

Twitter has sold its three-step process to many companies. However, we use it in our teaching as an example of flawed data-driven decision-making.

Comparing customers who saw a brand’s content with customers who did not see the content is like comparing apples and oranges. These customers differ in many other ways. Gwenn & Jenny’s most loyal customers are more likely to engage with the brand on Twitter, and they are also more likely to buy the brand. They don’t buy because they saw the brand appear in their Twitter feed. These loyal customers buy because they like the product, and because of this they also follow the brand on social media — not the other way around. Twitter’s approach dramatically exaggerates the impact of advertising on sales.

Data-driven decision-making empowers data providers and data scientists. The risk is that decision makers take data that is consistent with their preexisting beliefs at face value.

Moving to Decision-Driven Data Analytics

To move to a decision-driven data analytics approach, a company must start by identifying the business’s key decisions and the people who make them, and finding data for a purpose rather than finding a purpose for the data at hand. (See “Data-Driven Versus Decision-Driven Data Analytics.”)

DATA-DRIVEN VERSUS DECISION-DRIVEN DATA ANALYTICS

| Data-Driven Decision-Making | Decision-Driven Data Analytics |
|------------------------------------------|-----------------------------------------|
| Anchor on data that is available. | Anchor on a decision to be made. |
| Find a purpose for data. | Find data for a purpose. |
| Start from what is known. | Start from what is unknown. |
| Empower data scientists. | Empower decision makers. |

When we tell executives about the decision-driven approach, some are quick to point out a potential problem. They caution that decision makers who use data to support a decision that has already been made may be falling prey to confirmation bias. But that is not decision-driven data analytics. That is *preference-driven data analytics*, and it might be the worst way to make decisions — but it's unfortunately a very common one.

To make the shift to decision-driven data analytics and avoid the pitfalls of being preference-driven, leaders can take three important steps. First, it is the responsibility of decision makers to form a narrow consideration set of alternative courses of action. Second, it is the joint responsibility of decision makers and data scientists to identify the data needed to figure out which course of action is best. The third step is to choose the best course of action.

Step 1: Identify the alternative courses of action. Decision makers should start by thinking “wide then narrow.” Many decisions are made on autopilot, after only one course of action is considered. This can harm the quality of decision-making. Thinking *wide* means generating many alternative courses of action. To illustrate, let's go back to RollingBoulder. If the business objective is to increase the value of its customer base, a thank-you gift to reduce churn is just one of many possible courses of action. The company could also improve customer development (through up- or cross-selling, for example), make investments to acquire new customers (such as through sales promotions), or improve editorial content (perhaps by hiring new writers).

However, too many alternatives can make the problem intractable from a managerial and data-analytics perspective. Thinking *narrow* means that decision makers use their judgment to winnow courses of action. For instance, the customer relationship manager at RollingBoulder may realize that improving editorial content falls outside of her responsibilities.

By thinking “wide then narrow,” decision makers increase the likelihood that the final consideration set includes high-quality and feasible courses of action.

Step 2: Determine what data is needed in order to rank alternative courses of action. Decision makers and data scientists need to develop criteria to discriminate among and rank the alternative courses of action selected in step 1. The goal

of data analytics is to turn unknowns into knowns so that alternative courses of action can be ranked more objectively.

Starting from the decision draws attention to unknowns, and this has a major advantage. It makes it immediately obvious that there are limits to what can be known, and that unknowns can be tackled in many different ways. For instance, if you tell people that Interbrand has determined that the Mastercard brand is worth \$11 billion and ranks 57th in the world, most people will take this point estimate and rank at face value.⁵ If instead you ask people what they think the Mastercard brand is worth, most will realize that brand value can be quantified only imperfectly, and in a variety of ways. They are correct. Kantar Millward Brown estimates the Mastercard brand to be worth \$108 billion (ranked 10th in the world), and Brand Finance estimates it at \$19.8 billion (ranked 86th in the world).⁶ Starting from what is unknown highlights that the world is complex and uncertain.

Decision-driven data analytics is not about collecting as much data as possible. It's crucial to consider the value of data. If you would make the same decision before turning an unknown into a known as you would after, then there is no benefit in engaging in data collection and analysis.

Oftentimes, data collected for the purpose of making a decision has more value than data that's already available. RollingBoulder had to decide whether to add a gift to a customer's renewal letter. To make this decision, the company needed to know how sending a gift influences a customer's likelihood to churn. This question cannot be addressed based on the available data. It requires running a randomized controlled trial (an RCT, or A/B test): Customers are randomly selected to receive the gift or not, and the company then observes which customers churn and which customers stay.

Step 3: Select the best course of action. The final step should be easy. If the first two steps were executed well, data analytics will now reveal the best course of action.

Analyzing the data from the RCT, RollingBoulder learned an important lesson. The gift successfully reduced churn likelihood for some customers, but it backfired for others. The customers who were already least likely to churn without the intervention became even less likely to churn when targeted with the intervention. However, the customers who were

most likely to churn without the intervention became even more likely to churn when targeted with the intervention. In other words, the company had always been targeting precisely the wrong customers, as do many other businesses that blindly follow the “best practice” of targeting high-risk customers.

PABLO PICASSO ONCE SAID: “Computers are useless. They can only give you answers.” Decision-driven data analytics emphasizes the importance of asking questions and thus the importance of managerial judgment. This approach draws attention to unknowns and the value of additional data collection and analysis. Companies and leaders who take this approach benefit by ensuring that analytics initiatives are tied to action, are focused on answering the questions that matter, and challenge rather than bolster executives’ beliefs about how the world works.

REFERENCES

1. “Closing the Data Value Gap,” white paper, Accenture, Dublin, 2019.
2. RollingBoulder is modeled after a company described in E. Ascarza, “Retention Futility: Targeting High-Risk Customers Might Be Ineffective,” *Journal of Marketing Research* 55, no. 1 (February 2018): 80-98.
3. “Customer Attrition,” accessed Nov. 3, 2020, <https://en.wikipedia.org>.
4. “About Measuring Sales Impact,” Twitter, accessed Nov. 3, 2020, <https://help.twitter.com>.
5. “Best Global Brands,” Interbrand, accessed Nov. 3, 2020, <https://interbrand.com>.
6. “BrandZ Top 100 Most Valuable Global Brands 2020,” BrandZ (June 30, 2020); and “Global 500 2020,” Brand Finance (January 2020).

***Bart de Langhe** is a behavioral scientist and marketing professor at Esade – Ramon Llull University. **Stefano Puntoni** is a professor of marketing at the Rotterdam School of Management at Erasmus University and the director of the Psychology of AI lab at the Erasmus Centre for Data Analytics.*

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.

Empowering a Data Culture From the Inside Out

Transforming a company into a truly data-driven business involves fundamental organizational changes.

JONATHAN TUDOR, INTERVIEWED BY ALLY MACDONALD

For companies that struggle with data transformations, underthinking organizational change is often a bigger problem than technology issues. A company can have powerful tools and meaningful data at its disposal, but without the proper education and processes to put that data in the hands of the right people and provide business context, extracting value can prove difficult.

In 2016, Jonathan Tudor founded a self-service data program at GE Aviation aimed exactly at this problem. By recognizing that success would depend on empowering users beyond the data engineering and analytics teams, he was able to encourage buy-in from across the organization, increase engagement, and create cross-functional partnerships.

Ally MacDonald, senior editor at *MIT Sloan Management Review*, spoke with Tudor about his work with self-service data and organizational transformation. What follows is an edited and condensed version of their conversation.

***MIT Sloan Management Review*: What is a self-service data program?**

JONATHAN TUDOR: The idea with self-service data is, rather than hiring endless numbers of highly competitive data talent, why not take your existing intellectual capital and people capital within the company and empower them to do their own data analytics work? In a self-service system, line-of-business professionals and analysts in the company can access and work with data and data visualization directly, and they are supported by, but not dependent on, IT and data professionals to carry out their work.

This kind of program allows companies to remove technical boundaries and empowers people to use their own subject matter expertise — after all, they know the problems they're tackling best,



and they know what data they need — to generate insights and execute their work.

How has this new approach evolved to play a role in organizations?

TUDOR: It's helpful to look back in history. When we think of business intelligence [BI] and data warehousing, this has often been siloed within organizations. BI teams do all the work to gather the data, structure it (and hope it's structured correctly), and then deliver those insights to the customer. That process is similar to what we see in a waterfall method in IT, which is very sequential and dependent on the people carrying out each task.

But as companies push ahead with more data than ever and are utilizing things like data lakes, a waterfall approach becomes less effective. When data is fundamental to how you run your

organization, analytics are needed in all parts of the company very quickly and are key to business outcomes. The truth of the demands for many organizations is that they will never be able to hire enough data engineers or data scientists to meet these growing business needs on their own.

Self-service data allows companies to meet these demands by focusing less on who is carrying out the work, in favor of the business outcomes.

What KPIs are important when it comes to measuring the effectiveness of self-service data?

TUDOR: Companies will vary in their approach and certain metrics will be more important depending on the stakeholder. At GE Aviation, we focus on three self-service data metrics that shed light on financial benefits, how much value users are getting from the program, and innovation viability for the business.

First, there are financial KPIs. We can determine where there are impacts on the balance sheet income statement or how much revenue we can track from a return perspective.

The second important metric is utilization. We look at the number of unique users and track what percentage of them are staying active in a given period. For example, we might find that the self-service program has achieved 2,000 unique users across the ecosystem in the last three months.

A third area involves creating an innovation pipeline from tracking engagement and usage. For example, we look at every data artifact that has over 50 unique users in a month — every BI report, every analytic. Based off that, we can inform leadership about what is proving most important to users and may be worth further business investment. This helps us to create an organized pipeline that shows the value proposition of new, innovative ideas within the organization.

If self-service systems are key for enabling faster decisions in the organization, that means many teams and units throughout the business are now becoming more hands-on with the data. What does that mean for data governance?

TUDOR: Self-service data and governance definitely go hand in hand. What I liken it to is, you're trying to run a playground: You want to empower people so

that they can do what they need to do and do it well, but you also want to keep them from hurting themselves or others.

There's a delicate balance here from the start. You need to protect users and the company, ensure that you're compliant and meeting regulations, and enable better understanding of data. But at the same time, you don't want to be such a barrier that people can't do their work.

I'm a strong believer in building self-service programs from the ground up with governance and an understanding that they will need to scale. Things like data cataloging, data lineage, and providing the business context around the data are all very important parts of governance, because people need to have the right information and context about the data they're looking at in order to be successful.

What are other challenging areas that companies should be mindful of when approaching these types of initiatives?

TUDOR: Data stewardship is also very important. Today there are lots of ways for companies to automate their data catalog using software to help make decisions. But ultimately, you still need to have people in the company acting as data stewards — that is, providing business context and serving as points of contact for others in the organization who may have questions or need more information. This is often very difficult to achieve in companies, because being a data steward is not commonly a dedicated role, and it's challenging to get people to commit their time to something that is not necessarily their primary job.

Something we learned early on to help incentivize data stewardship involved implementing gamification. We started tracking activity of users in our ecosystem and created a related point system and leaderboard that ranked people using the self-service program. People could then compete against one another and be recognized monthly for their contributions.

I've seen many organizations struggle with this, and I will admit this is still hard today. But by making it more interactive and even fun, we've significantly improved participation.

What should data-focused managers be thinking about when it comes to onboarding talent today? What do companies and managers overlook?

TUDOR: The talent you're looking for often needs to be different than [what] we generally think about when we talk about IT or data and analytics. You can't just be a technologist today. Oftentimes, the problem you're going to encounter won't be a technology problem but a people problem.

On my team, we often think of ourselves as a startup within a larger organization, where you need to wear many different hats. We have many people who have nontraditional backgrounds for IT — for instance, former musicians and chefs who switched careers and bring diverse experience to their work in data and technology. I think this diversity helps in many ways for connecting with people, because much of our work is people-oriented *in addition* to being driven by good technology.

The other thing that's really important for training talent is stressing the importance of the business partnership that needs to take shape for any self-service program to be successful. For example, we train and partner with dozens of data ambassadors — professionals not on our central data team but who can act as extensions of the team in locations across the globe. They help manage communication, break down organizational barriers, and drive governance across different sites. Encouraging this type of collaboration is critical for amplifying the role of a smaller self-service team across the wider organization.

It's clear from our discussion that these systems aren't static — teams are continuously measuring and learning. How do managers shift their thinking and way of managing traditional business intelligence to support self-service data success?

TUDOR: One thing for managers is simply being OK with data and analytics or IT not being done by teams who are in data and analytics or IT. That's a very fundamental one, but it can be very hard. You have to be comfortable with letting go of some control.

In some ways, it's like you're an app store and you're helping a self-service community build apps that might not be developed on your team. This means becoming more of a systems thinker in order to drive a design that empowers other people to be successful at scale. This is likely the biggest and most difficult mind shift to make, but there's a lot to gain from embracing it.

Jonathan Tudor is a director of data and analytics at GE Aviation. He has been a part of the data lake at GE Aviation since its inception and has led strategy and initiatives for data ingestion, engineering, architecture, self-service, analytic invention, automation, and governance. Ally MacDonald (@allymacdonald) is senior editor at MIT Sloan Management Review.

Copyright © Massachusetts Institute of Technology, 2020.

All rights reserved.



As an enterprise strategist at AWS, **Ishit Vachhrajani** helps organizations increase speed and agility, drive innovation, and create cloud-based operating models to better serve customers. Previously, he was chief technology officer at A+E Networks, where he oversaw global technology and served as the company's head of global products, applications, and data. He also held technology leadership positions at NBC-Universal and various consulting organizations. He received a bachelor's degree in instrumentation and control engineering, with a gold medal for academic achievement, from the Nirma Institute of Technology in India and graduated from the Executive Leadership Development Program at the University of Virginia's Darden School of Business.



Developing the Data-Driven Organization: Leadership, Culture, and Learning

In this Q&A, Ishit Vachhrajani, enterprise strategist at Amazon Web Services (AWS), describes the cultural and organizational changes that enterprises must make to become genuinely data-driven. He discusses how to establish key leadership roles, address governance, overcome resistance, and generate widespread excitement about and support for AI issues.

This conversation has been edited for clarity, length, and editorial style.

Q: What does it mean to be data-driven, beyond just the technology and process considerations?

Vachhrajani: Being a data-driven organization means culturally treating data as a strategic asset and then building capabilities to put that asset to use not just for big decisions but also for everyday action on the front line. Organizations that are successful in this way focus on capturing, cleaning, and curating meaningful data, making that data available widely across the organization, having the skill set and proficiencies to convert that data into actionable insights, and then having the culture that allows people to take action on those insights.

“Job one for data leaders is really driving education around data. What that means is: Demystify it. Make it relatable so that it is something the whole organization can experiment with, touch, feel, and use on a daily basis, and then directly tie it to business outcomes.”

Q: What kind of top leadership is important for building a data-driven culture?

Vachhrajani: Having a single-threaded leader helps a lot, especially when you are starting out. The way we define a single-threaded leader at Amazon is someone who wakes up and goes to sleep thinking about just this one thing. This could be the chief data officer; it could be the chief analytics officer; it could be any other senior leader.

I often advise customers to pick someone who is senior but also well respected within the business, because this change cannot simply be driven by authority. It has to be driven by influence. Having someone who is well respected within the business, but is also senior enough, really helps drive these initiatives forward. You may then decide to decentralize and make sure that these capabilities are actually built in the rest of the organization. The idea of the single-threaded leader is not to build an empire of data, but to enable and bring the rest of the company together.

Q: How about data leaders? What role should they play in creating and setting priorities for this type of culture?

Vachhrajani: Job one for data leaders is really driving education around data. What that means is: Demystify it. Make it relatable so that it is something that the whole organization can experiment with, touch, feel, and use on a daily basis, and then directly tie it to business outcomes. Next is building the capabilities to allow data leaders to do that — the technology, the tools, and the right governance. Third is removing the impediments, working with other internal stakeholders, to make sure that information and data flow freely within the company.

Also, dealing with resistance is a big part of becoming data-driven. Resistance is real because while data is very, very empowering, it can also evoke strong emotion. A lot of times, people use data to justify what has already happened or been decided upon, rather than to actually guide them. There's also a bit of a fear about losing control of the narrative, the story, when it comes to democratizing data.

One big way to address that is by communicating the intent very clearly and strongly. This is similar to driving a lot of other types of changes, not just with data: You announce the intent from the top, then show action that ties to that intent.

This also brings up the issue of access. There are valid reasons why some people shouldn't have access to data: compliance, privacy, and so on. But becoming data-driven also involves changing those gatekeeping functions to be more enabling, to treat data not as a departmental property, but truly as an organizational asset. In

other words: Departments don't own the data; they steward the data.

“Dealing with resistance is a big part of becoming data-driven. Resistance is real because while data is very, very empowering, it can also evoke strong emotion. A lot of times, people use data to justify what has already happened or been decided upon, rather than to actually guide them.”

So how do you really change that gatekeeping role? The job of, let's say, a finance or marketing or sales or operations or production person shouldn't be that they own data and they decide who gets access to it. Instead, because they deeply understand that data, they take on the role of educators. They take on the role of publishing, cleaning, and curating meaningful data, and then making it easier for everybody else to consume.

Q: What do companies need to know about AI governance?

Vachhrajani: Governance is a very important topic. The way I think about governance is: Define your clear tenets, and your governance framework will follow. Before we talk about controls, what are the underlying guiding principles and tenets? This can be as simple as “Our goal is to enable more access to data, not to restrict it.”

It's also important to ensure that data lineage is factored in from the get-go so that you aren't implementing remedial controls that are reactive in nature when regulations change or when you have access issues. You are factoring in that data lineage all the way from where data originates to how it is used and consumed, and mapping that process out. You're implementing privacy, security, and compliance considerations at every step of that process, by design.

Another issue is having a mechanism for data validation. I'm not just talking about technical validation; I'm also talking about business validation of data. That's because often data may be mathematically correct, but when you apply the business logic, it doesn't make sense. So it's important to automate and integrate those

aspects of validation in the governance process so that when the data is published, it is stamped not just from a basic quality check but also from a business validation standpoint for the consumer.

“Most organizations, and most people, learn by doing. So I recommend picking one or two high-impact use cases, where there is a strong hypothesis that can be built, where you think data and analytics will actually make a difference in achieving a business outcome. Then run those use cases through this new approach that you’re taking, and communicate about that broadly.”

Finally, from a technical aspect, we always advise customers to encrypt everything. Encrypt all your data at rest. Encrypt data in motion. These are simple things, but putting them down as defining tenets can provide teams, especially smaller autonomous teams, a framework within which they can operate.

Q: What about some tactical actions that companies can take to build interest and support for these advanced capabilities throughout their organizations?

Vachhrajani: I think it comes down to the fact that most organizations, and most people, learn by doing. So I recommend picking one or two high-impact use cases, where there is a strong hypothesis that can be built, where you think data and analytics will actually make a difference in achieving a business outcome. Then run those use cases through this new approach that you’re taking, and communicate about that broadly. Even if those use cases may not be applicable to the whole organization, it’s important to communicate about them widely to build excitement, to build momentum.

We hear a lot about data literacy. I actually prefer the term *data proficiency*. That’s because it’s not just about being aware of the data; it’s about how to create a higher level of understanding and processes within the organization to put the data to use. Companies can build their data proficiency from the ground up by using creative ways to find education options.

Supporting a Data-Driven Visual Discovery Engine – and a Unique User Base

David Chaiken, Pinterest’s chief architect, can sum up his company’s data strategy in one sentence: “Pinterest is data-driven, top to bottom, in everything we do — but we never lose our focus on the individual people who use our product.”

Founded in 2010, Pinterest describes itself as a “visual discovery engine more than 400 million people use each month to find inspiration for their lives.” Users, known as Pinners, have saved more than 240 billion “pins” — information on favorite topics ranging from recipes and travel to crafts, home, and style ideas — which other Pinners can find through searches and recommendations. “We come to work every day trying to figure out how to provide value to all those people,” Chaiken says.

Technology, of course, plays a major role in supporting that mission. “Almost all changes to our product are fundamentally data-based and come from the research we do on our users,” he says.

Using data and machine learning, the company collects and analyzes, on a millisecond-by-millisecond basis, trending searches and overlapping interests across the community to identify relevant ideas for each person. “Then there’s the ‘why’ — analyzing what we should do and analyzing user response to changes we make,” Chaiken says.

Pinterest also constantly tracks not just how, but where, it can best serve users. “For instance, we started out as a website, but fairly quickly realized that we needed to be on mobile phones, and now 70% of all Pinners use Pinterest on mobile,” he says.

Meanwhile, Amazon Web Services’ infrastructure-as-a-service capabilities have contributed significantly to Pinterest’s remarkable growth. “AWS has provided us with enormous value by being able to scale and providing the reliability we need to run our business,” Chaiken says. “And as a born-in-the-cloud company, our ambitions have never been limited by the walls of a data center. Working with AWS has removed that whole set of constraints.”

For example, I've tried going beyond training and certification programs to offer data hackathons, where we would take publicly available data sets that everyone can relate to in their day-to-day lives, such as parking tickets, traffic violations, census data, and restaurant bookings. Then we bring in nontechnical folks from the front line to play and model with that data. We'll provide them help and resources, then give them fun problems to solve. For instance: "Find me a restaurant in Manhattan on a Friday night where I have the lowest risk of getting a parking ticket." This allows them to use data from everyday life to model multiple different scenarios, build visualization, build models, and combine data sets, which can then ultimately translate into actions within the business.

Q: Is there anything else that you'd like to add?

Vachhrajani: Talking about this brought one more point to the fore. We often talk about consuming data. When we're thinking about becoming data-driven, most of the discussion is around finding data, bringing in more data, running analytics, things like that. It's also important to have a very strong product and application strategy that ties into capturing data, because, especially in business, a lot of transactions still happen on the back of the napkins, in spreadsheets, from people talking to people.

One thing I do often is just observe the flow of information in the organization. Who is asking the question, who is providing the information, and who is making the decision? If you look at those three aspects, there are a lot of opportunities, before you even start to figure out how to consume that data, for bringing those lost insights into the workflow. It's not just data as a byproduct of everything else. You should treat data as a product, something that you manage end to end throughout the life cycle.

Another aspect goes back to the resistance issue. Often, when companies start on this journey, the option to use the old way of doing things is kept open. If you do that, inertia will always pull you back to that old way. If you want to drive change, you've got to remove the option to go back. Of course, you have to do it in a mindful, intentional fashion, making sure things are working and running well. But ultimately, to succeed, you have to eliminate the options for going back and doing things the way you've always done them.

ABOUT AWS

For 14 years, Amazon Web Services has been the world's most comprehensive and broadly adopted cloud platform. AWS offers more than 175 fully featured services for compute, storage, databases, networking, analytics, robotics, machine learning and artificial intelligence, internet of things, mobile, security, hybrid, virtual and augmented reality, media, and application development, deployment, and management. Services are offered from 77 Availability Zones (AZs) within 24 geographic regions, with announced plans for nine more AZs and three more AWS Regions in Indonesia, Japan, and Spain. Millions of customers — including the fastest-growing startups, largest enterprises, and leading government agencies — trust AWS to power their infrastructure, become more agile, and lower costs.

Learn more at aws.com/big-data or aws.com/databases.



PDFs ■ Reprints ■ Permission to Copy ■ Back Issues

Articles published in *MIT Sloan Management Review* are copyrighted by the Massachusetts Institute of Technology unless otherwise specified.

MIT Sloan Management Review articles, permissions, and back issues can be purchased on our website, **shop.sloanreview.mit.edu**, or you may order through our Business Service Center (9 a.m. - 5 p.m. ET) at the phone number listed below.

Reproducing or distributing one or more *MIT Sloan Management Review* articles **requires written permission.**

To request permission, use our website **shop.sloanreview.mit.edu/store/faq**, email **smr-help@mit.edu**, or call 617-253-7170.