CRITICAL SUCCESS FACTORS TO ACHIEVE A BETTER ENTERPRISE DATA STRATEGY IN A MULTI-CLOUD ENVIRONMENT

As technology has become more powerful, more affordable, and more accessible, we are witnessing new breakthroughs and discoveries at a pace that was previously unimaginable. At Cloudera, we are inspired by our customers who harness the power of data to solve important problems, improve their customers' experience, make their businesses run more smoothly and profitably, and, in many cases, literally save people's lives.

At the same time, we've seen new struggles emerge that threaten to interrupt the clip of innovation if not grind it to a halt. Enterprise IT is simply not meeting business demands, and users are prioritizing speed over security, accuracy, and maximum impact. A few points to consider follow.

Data, the lifeblood of all business, is pouring in from sensors, networks, applications, and an ever-expanding horde of connected devices. But do enterprises even have access to all their data or are they blind to the data that matters most? Are the teams that need that data waiting in an endless queue or are they creating their own "shadow data" copies and working from the best they can cobble together? Are they able to properly secure and control the data regardless of where it lives? Can they do this efficiently and at scale?

Next, companies increasingly operate in a complex web of public, hybrid, and on-premises environments. How are they managing this new ecosystem to ensure optimal data use, portability, compliance, and security? What is the plan for continued cloud migration? How are companies thinking about or managing the patchwork of multi-vendor and proprietary, legacy, and modern infrastructures, such as the Edge? Are they feeling locked in or clamoring to avoid that fate?

Finally, companies have never had more options to work with their data, from analytics to artificial intelligence. Unfortunately, many of these functions are limited to one cloud, one data set, and one stream. What if companies could unlock all their analytic options for all their data? What would it take to break down the silos and bring in the right function at the right time regardless of where the data resides? Do companies have a data strategy that's working? If not, what's standing in the way?

Cloudera sponsored research by Harvard Business Review Analytic Services to get answers to these and many more questions. Beyond our own experience, we wanted to understand the aspirations and challenges in data management; the pain points in cloud strategy and cloud migration; the compromises and blind spots organizations make and have about analytic functions; and the unique challenges around security, privacy, and governance in today's hacker-happy and increasingly regulated environment. The results were astonishing.

For the vast majority of executives, managing and using data resources is important for creating business value. However, over 40% don't feel they have a clear strategy for managing data across multiple clouds. Furthermore, over 60% feel their organization does not yet know what data they'll need in order to take full advantage of emerging technologies like AI, intelligent automation, and augmented reality, which have the potential to fuel the next era of their data transformation.

One thing is clear: Organizations need a new way to think about data, and data must be the driver. Starting with one cloud or one function and trying to apply it to the whole data strategy is inviting limitations and accepting a weaker outcome. Enterprises want to say yes to multi-cloud and to storing and computing their data on the most optimal infrastructure without limiting accessibility. Enterprises want to say yes to all the analytic functions their business requires. They must be able to say yes to security without limitations. And, to prevent vendor lock-in and ensure the greatest opportunity for innovation and access, enterprises want to say yes to open source, open APIs, and open compute. This report reveals the strong need for an enterprise data cloud—a new category of analytics and data management that helps enterprises derive value from data across any environment, run multi-function analytics on any data from the Edge to AI, and secure and govern it all—wherever it lives.

At Cloudera, we believe that data can make what is impossible today possible tomorrow. We're going to show you how to get there.

—ARUN C. MURTHY, CHIEF PRODUCT OFFICER, CLOUDERA

CRITICAL SUCCESS FACTORS TO ACHIEVE A BETTER ENTERPRISE DATA STRATEGY IN A MULTI-CLOUD ENVIRONMENT

Data may be the fuel that powers the modern enterprise, but it is only as useful as the ability to explore, analyze, and transform a deep well of information into intelligence and action. The most exciting organizations today bring an insight-driven approach to everything they do—from evolving their business models to embedding analytics capabilities across the organization and incentivizing corporate culture to become more galvanized by data and the interpretation of the story it tells.

To achieve the promise of an enterprise data strategy, most organizations plan to increase dramatically their use of analytics and cognitive capabilities for everything from business intelligence (BI) to machine learning to artificial intelligence (AI). A Harvard Business Review Analytic Services survey of 185 global executives across a wide range of industries finds that the vast majority (73%) believe that their data resources hold the key to creating business value—including 33% who give it the highest possible rating: 10 out of 10 in terms of importance. More than half (51%) plan to leverage multiple cloud providers to achieve this strategy.

"Our data platform is like electricity," says Mike Flynn, principal at PwC. "We want our data platform to be super easy and super fast for all of our teams around the world to plug into it."

The survey finds, however, that there are a number of challenges to realizing the full potential value of enterprise data that may be distributed across private, hybrid, and multiple public-cloud environments. Data siloes, data lock-in, and lack of interoperability are among the greatest strategic impediments that organizations face, according to respondents—all problems that can intensify without a strategy to manage data spread across multiple clouds.

At the same time, security and governance are also tightening as more jurisdictions—and even organizations themselves—place more restrictions on who can access which data, and where and how they can access it. Tracking data lineage—who touched what data and where and what they did to it—is already a challenge. Tracking that data while it's migrating from place to place or proliferating into multiple copies across multiple cloud environments adds to the complexity.

HIGHLIGHTS

51%
OF RESPONDENTS PLAN TO LEVERAGE MULTIPLE CLOUD PROVIDERS AS PART OF THEIR DATA STRATEGY, BUT ONLY 34% HAVE AN EFFECTIVE STRATEGY FOR MANAGING DATA IN THE CLOUD.

49%
PLAN TO BUILD MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE CAPABILITIES OVER THE NEXT THREE YEARS.
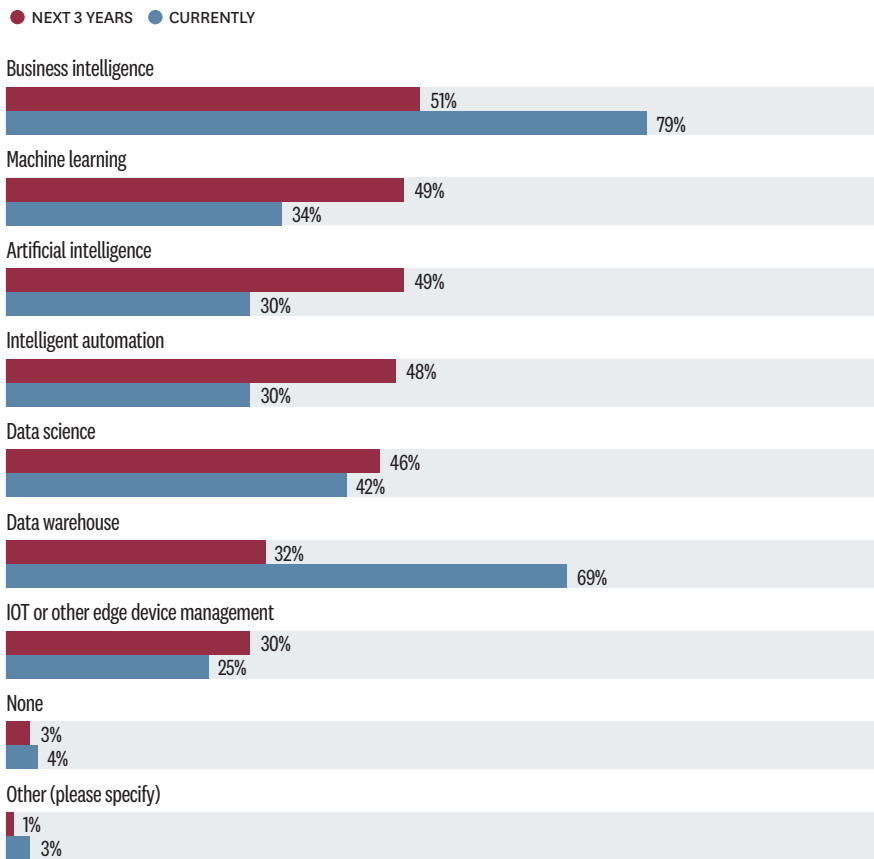
69%
SAY THEIR ORGANIZATION NEEDS A COMPREHENSIVE DATA STRATEGY IN ORDER TO MEET ITS STRATEGIC GOALS OVER THE NEXT THREE YEARS, YET ONLY 35% SAY THEIR ORGANIZATION'S ANALYTICS AND DATA MANAGEMENT CAPABILITIES ARE ON COURSE TO MEET THOSE GOALS.

> "By 2022, more than half of major new business systems will incorporate continuous intelligence that uses real-time context data to improve decisions… It's a grand challenge—and a grand opportunity—for analytics and business intelligence teams to help businesses make smarter real-time decisions in 2019."
> —Gartner[1]

THE FUTURE OF DATA ANALYTICS: INTELLIGENT FUNCTIONS

Data analytics organizations plan to add/enhance over the next three years versus the type of data analytics organizations currently engage

● NEXT 3 YEARS ● CURRENTLY

Business intelligence
51%
79%

Machine learning
49%
34%

Artificial intelligence
49%
30%

Intelligent automation
48%
30%

Data science
46%
42%

Data warehouse
32%
69%

IOT or other edge device management
30%
25%

None
3%
4%

Other (please specify)
1%
3%

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

This paper will examine several key interrelated themes that define the enterprise data landscape today, starting with how organizations must engage multiple analytic frameworks to gain intelligence and insights. This aspect of the task involves managing data wherever it is: on-premises, in the cloud, or in multiple cloud environments. Complicating this situation is the growing complexity of governance and security, as well as the importance of data mobility and accessibility. To achieve all this, of course, the organization needs a comprehensive data strategy that works in a multi-cloud environment.

Engaging Multiple Analytic Frameworks

Delivering insights when and where they are needed requires identifying the right data, ingesting and transforming it, and determining the right analytical framework to make sense of the data for each purpose. The most transformative use cases today— think autonomous vehicles, augmented reality, and individualized medicine— are built on a range of data science tools, analytic frameworks, machine learning, and AI.

In fact, machine learning, AI, and intelligent automation will play a far larger role in data analytics over the next three years, respondents say. Half of the executives surveyed plan to add AI capabilities, compared with 30% who have them now. FIGURE 1

Gartner predicts that by 2022, more than half of major new business systems will incorporate continuous intelligence that uses real-time context data to improve decisions. Continuous intelligence integrates

real-time analytics within a business operation, processing current and historical data to prescribe actions in response to events. "It provides decision automation or decision support. Continuous intelligence leverages multiple technologies such as augmented analytics, event stream processing, optimization, business rule management and ML [machine learning]."[2]

"Continuous intelligence represents a major change in the job of the data and analytics team. It's a grand challenge—and a grand opportunity—for analytics and business intelligence teams to help businesses make smarter real-time decisions in 2019. It could be seen as the ultimate in operational BI [business intelligence]."[3]

At the same time, stream and edge processing will become more important—particularly as the 5G mobile network removes barriers to continuous communication with edge devices everywhere. Companies across a range of industries will be able to track what's happening at the edges of their businesses—and they will need to manage these new workloads in a cost-effective way.

Logistics companies, for example, will be able to stream data in real time to the engines of their trucks or monitor temperature in real time as a container moves from a ship to a train to a warehouse. This will require a range of analytic functions, including real-time data streaming from sensors on a vehicle or container, along with the ability to store, analyze, and apply machine learning algorithms or AI that can deliver functionality back to the edge devices to make intelligent decisions.

But the survey reveals that most organizations still have a long way to go before they realize anything like a continuous-intelligence business model. Only 40% of executives say they have found a way to move data into the right environment, a mere 22% can ingest data fast enough to make decisions in real time, and just 20% have developed stream processing capabilities. FIGURE 2

There is another factor in developing any analytic framework: The quality of the insights and intelligence gained will always come back to the quality of the data. "The hardest part about AI is getting good annotated data sets on relevant topics," says PwC's Flynn. "For much of what we're trying to automate, getting an annotated data set means that I have to get a subject matter expert in a specific area to go in and tell me what's important and to do it in a controlled and repeatable manner, so that I can start to train the machine. Those data sets are going to become the value of what we're doing, because the AI algorithms, for the most part, are public."

Managing a Multi-cloud Environment

For most organizations, information is created, collected, moved, and stored across multiple locations, in multiple clouds. Sensitive data might be kept on-premises, in a private cloud, or under tight access restrictions in a public cloud. Different lines of business may be running their workloads through specialized cloud service providers. Some of the data with the most promising insights in a specific field may be owned and managed by a third party that has curated a data store that is more comprehensive than what any one organization could assemble with only proprietary data. Likewise,
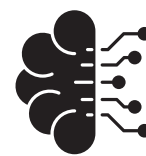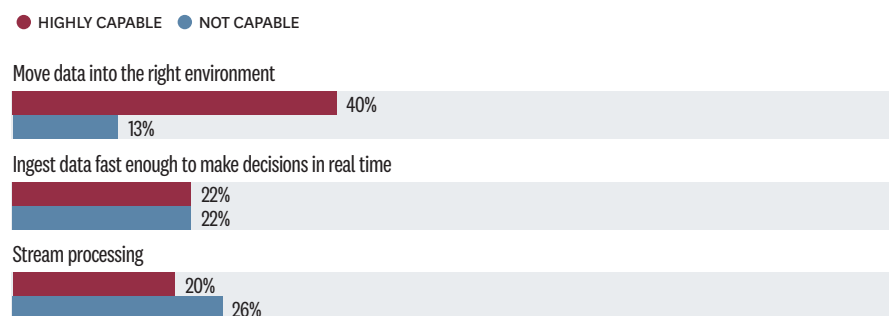
HALF OF THE EXECUTIVES SURVEYED PLAN TO ADD AI CAPABILITIES

FIGURE 2

MOST DATA STILL MOVES IN THE SLOW LANE

Respondents rating their organization's ability to perform each of the following functions

● HIGHLY CAPABLE ● NOT CAPABLE

Move data into the right environment
40%
13%

Ingest data fast enough to make decisions in real time
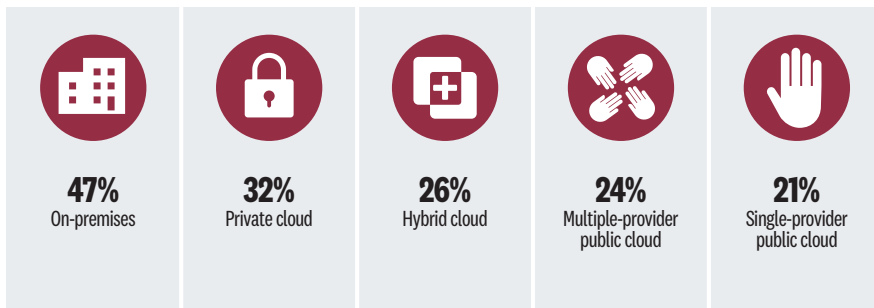22%
22%

Stream processing
20%
26%

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

FIGURE 3

DATA MANAGEMENT IS SPREAD ALL OVER

Where respondents' organizations manage data

47%
On-premises

32%
Private cloud

26%
Hybrid cloud

24%
Multiple-provider public cloud

21%
Single-provider public cloud

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

FIGURE 4

NOTHING CLEAR-CUT WITH MULTIPLE CLOUD CLIMATES

Organizations having a clear strategy for managing data across multiple cloud environments

15% DON'T KNOW

41% NO

44% YES
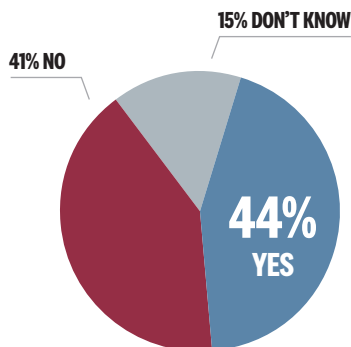
SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

the best path to an AI solution might go through a third-party cloud service.

"A hybrid or multi-cloud environment is not just for data surge and processing power anymore—it's for applications and analytics as well," says Thomas Davenport, the president's distinguished professor of information technology and management at Babson College. Indeed, cloud providers are making available certain types of applications for analytics or specialized data warehouses in the cloud. "If you have a preference, say, for a certain type of algorithm or analytical software or storage approach, then you may be more likely to find it in one cloud provider than in another," he explains.

Hybrid and multi-cloud architectures offer great opportunities and better control, putting organizations, business units, and even individuals in charge of deciding which analytic workloads to run, and where and when to run them. This presents a twin challenge: how to bring together the best data resources with the finest analytic tools when they are spread across multiple clouds— and keep track of who is doing what with the data.

While only 24% of organizations use multiple public cloud providers today, that number will rise. FIGURE 3 Half (51%) say they will need to leverage multiple cloud providers as part of their data strategy.

Nearly every organization desires the ability to analyze multiple diverse

data sets in real time, but few can. For example, Rick Spitler, co-CEO of financial IT management firm Novantas, says he sees "extraordinary value in the insights available by considering bank branch decisions in light of customer behaviors, in light of competitive price positions, in light of marketing spend." But to achieve that level of insight means drawing on disparate data sets over time. "We have all those different domains of data, but we have a massive integration issue with our cloud service as well as our own internal cloud," he observes.

Gartner recently warned that "data and analytics leaders must prepare for the complexities of multi-cloud and intercloud deployments, to avoid potential performance issues associated with data latency, unplanned cost overruns and the ongoing difficulties associated with integration efforts, data transfer fees and increased complexity."[4]

"Do not underestimate the added complexity of running multiple cloud environments; a single set of policies may need to be implemented multiple times with differing capabilities and tools," Gartner says in a recent report.[5] The report also states that "data and analytics leaders must accept this new reality, but can also create policies that mitigate risk and complexity."[6]

The survey shows that for most organizations, the opportunity to leverage new and better storage options, workload management, and specialized data sets and analytic tools from a variety of cloud providers clearly outweighs the added complexity and compromises of managing a multi-cloud environment. Yet a multi-cloud environment requires a multi-cloud strategy that works across all enterprise data in an open environment and with a coherent data governance policy—a strategy that only 44% of organizations have worked out. FIGURE 4

Establishing Trust: Security, Governance, and Regulations

Managing the availability, usability, integrity, and security of data used

"DO NOT UNDERESTIMATE THE ADDED COMPLEXITY OF RUNNING MULTIPLE CLOUD ENVIRONMENTS; A SINGLE SET OF POLICIES MAY NEED TO BE IMPLEMENTED MULTIPLE TIMES WITH DIFFERING CAPABILITIES AND TOOLS." GARTNER[7]

Finding the Right Cloud for Every Purpose

Most organizations see a future that includes the public cloud. "If they're thinking about doing anything with data analytics, they're thinking about how to do it in a hybrid cloud, if not in the public cloud completely," says Michael Franklin, the Liew Family Chair of Computer Science and chairman of the department of computer science at the University of Chicago.

Most respondents, however, have yet to move a substantial portion of their data or workloads to the public cloud. Nearly half (48%) manage less than 25% of their data in the public cloud. FIGURE 5

Problems with legacy applications (49%), security (42%), and cost (36%) remain the biggest issues with cloud service providers—and are the reasons that organizations are not making greater use of the public cloud today. Says Flynn: "When you start moving into the cloud, you're adding a middleman and you're paying a markup. If you don't actually optimize your workloads for the cloud and just try to replicate in the cloud what you do on-premises, you may end up paying more." In fact, some companies end up bringing their data back to on-premises environments to avoid mounting costs from poorly executed cloud strategies. FIGURE 6

PwC is taking a more strategic approach as part of the global partnership's digital strategy. "We want all of our workloads to be cloud-native or cloud-ready," says Flynn. "We want all of our capabilities to be ephemeral, meaning we can spin them up and spin them down when they're needed or not needed. That ability is what's going to drive the cost savings of being able to operate at scale globally, but to be able to do it cheaper than if I had all these different data centers spread around the world. I have to be able to turn off storage and programs when we aren't using them."

Seasonal surges in workloads are a case in point. For example, one PwC project involves looking for fraud and anomalies in journal entries by pulling ERP data and running analytics on it, the bulk of which happens over a six-month period every year. Two years ago, that process cost $1.6 million to run, but PwC has been able to drive the cost down to about $300,000 a year.
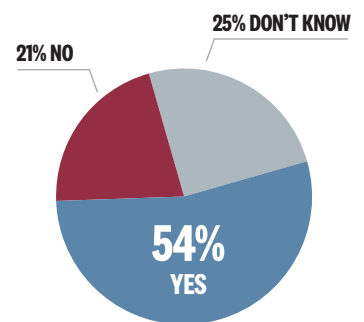
FIGURE 5

KEEPING IT PRIVATE: THE PUBLIC CLOUD IS NOT FOR EVERYONE, YET... MIGRATION TO THE PUBLIC CLOUD WILL EXPAND

Percentage of data managed in the public cloud
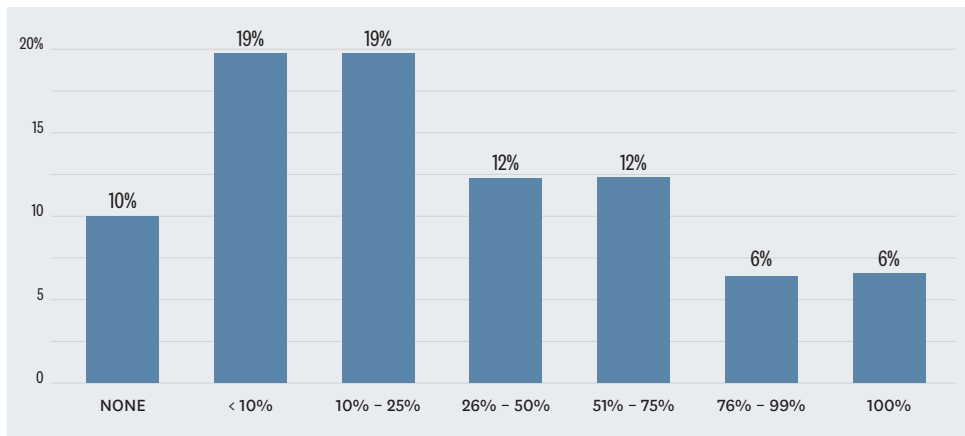
Plans to increase the amount of data it manages in the public cloud over the next 12 months

21% NO · 25% DON'T KNOW · 54% YES

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

> "We want all of our capabilities to be ephemeral, **meaning we can spin them up and spin them down** when they're needed or not needed," says Mike Flynn, principal at PwC.

across the enterprise has never been more important or more complex. Organizations have traditionally had to secure their data and control access to it. Now, they need to orchestrate governance and security across different infrastructure and storage systems. It is a question not only of adhering to the growing range of regulatory requirements but also establishing a broad governance policy that encompasses the ways that people, processes, and different technologies work with data in a compliant, auditable, and secure way.
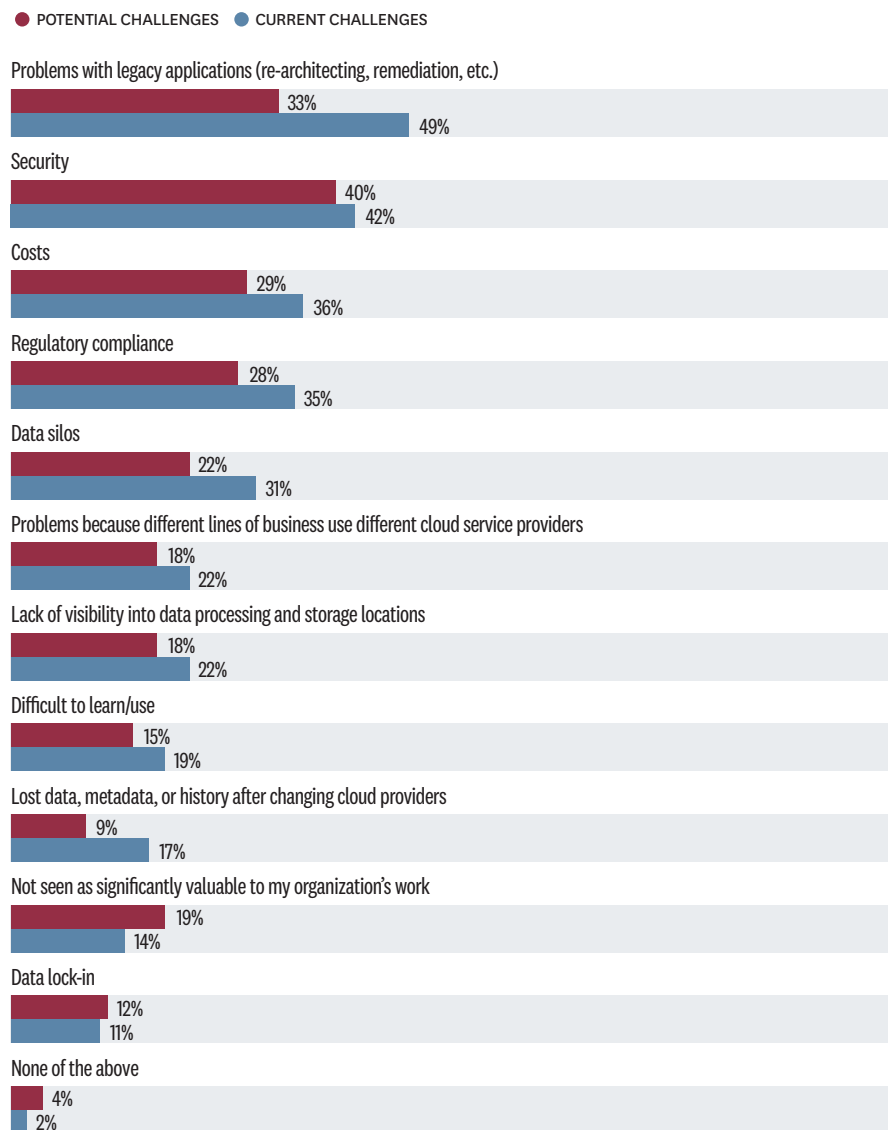
First, regulations around data movement and privacy will play an ever-larger role in analytic and data management strategies. More than three-quarters of respondents (77%) are required to secure data within a regulatory framework, and most (57%) expect to face new data privacy regulations in the near future. GDPR is the most commonly cited data regulation, affecting 61% of respondents.

Governance is not merely a regulatory issue. Clients, boards, and other stakeholders are imposing new requirements as well. With so much attention focused on data privacy and governance at a populist level, a lot of organizations are trying to prove they can regulate themselves. For Novantas, with global banking clients subject to their own geographic as well as financial industry regulations, navigating the web of regulations is already a challenge. "Some banks just view their data as so proprietary that they don't want it to leave," says Spitler at Novantas. But now, some clients are imposing additional restrictions on the nationality of anyone who touches their data, he says. "It's not just that the data can't leave the country or the

FIGURE 6

SECURITY COSTS REPRESENT BIGGEST CLOUD SERVICES CHALLENGES AHEAD

Challenges respondents have encountered with cloud services to date versus potential challenges that have held back their organizations from making greater use of the public cloud

● POTENTIAL CHALLENGES   ● CURRENT CHALLENGES

Problems with legacy applications (re-architecting, remediation, etc.)
33%
49%

Security
40%
42%

Costs
29%
36%

Regulatory compliance
28%
35%

Data silos
22%
31%

Problems because different lines of business use different cloud service providers
18%
22%

Lack of visibility into data processing and storage locations
18%
22%

Difficult to learn/use
15%
19%

Lost data, metadata, or history after changing cloud providers
9%
17%

Not seen as significantly valuable to my organization's work
19%
14%

Data lock-in
12%
11%

None of the above
4%
2%

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

IN THE SURVEY, ROUGHLY ONE IN EIGHT ORGANIZATIONS SAID THAT THE RISK OF VENDOR LOCK-IN HAS PREVENTED THEM FROM MAKING GREATER USE OF THE PUBLIC CLOUD.

bank—we can't even allow any non-citizen to access the data," he explains. Novantas has had to pull Canadian staff off of work on some U.S. banks, for example, and to stop using offshore developers in some cases.

There are governance priorities that go beyond the growing web of regional, industry, and company regulations. The ability to trace the provenance of information, maintain data integrity, and preserve all related metadata is essential for establishing trust in data that may be used for machine learning and other analytics that yield intelligence, train systems, or control actions. Tracking data provenance means there is a clear record of where it came from, what's happened to it over its lifetime, who has touched it—all of which should indicate how much the data can be trusted for specific use cases.

"As data flows through a process, or through an organization, or through a group of organizations, it gets modified, it gets summarized, it gets augmented," explains Michael Franklin, the Liew Family Chair of Computer Science and chairman of the department of computer science at the University of Chicago. "How do you track that so that when you're trying to make a decision based on data, you have some hope of understanding where that data came from? Companies are not able to get the full value out of their data because it's too dangerous to let it flow around. That's a lost opportunity because there is knowledge out there that doesn't exist in any one data set. You need to put data together to be able to extract that knowledge."

Building an Open Platform
To build a platform that can draw on the deepest well of information requires access to the right data in the right operating environment, through open compute architectures, open data stores, and open partner ecosystems.

The biggest risk to the open platform is vendor lock-in. "There is a conflict of interest where cloud providers want you to set up shop in their cloud and not leave," says Franklin. "This adds

huge risk to customers because they can find themselves at the mercy of a particular cloud vendor. The ability to move across vendors or to use multiple vendors at the same time is absolutely crucial." In the survey, roughly one in eight organizations said that the risk of vendor lock-in has prevented them from making greater use of the public cloud. FIGURE 6

At many organizations the goal is to be cloud agnostic, even if that is not always possible or practical. "You should be able to select the type of workload, select the capacity, run the job, and meet security requirements," says Kaushik Deka, a partner and CTO at Novantas. "Whether it runs on-premises or in a public or private cloud infrastructure, it should be agnostic to the use case. But that's not the case today." Deka would like to see an integrated data platform that would allow him to focus on the use case, the workflow type, and the security requirements—and not have to devote so many resources to moving data from one spot to another.

The view of what it means to be open and agnostic is evolving. At PwC, for example, the goal was to stay completely focused on open capabilities. "We did not want to rely on any platform-as-a-service, to avoid vendor lock-in," explains Flynn. PwC is still committed to open compute and storage, but the firm is also using more databases that are offered across a number of major cloud vendors. "We can update the scripts and spin up that database in any number of environments, and still maintain being cloud agnostic," he explains.

According to Gartner, "multi-cloud ISV offerings running on cloud infrastructure promise cloud

"Companies are **not able to get the full value out of their data** because it's too dangerous to let it flow around. That's a **lost opportunity**," says Michael Franklin at the University of Chicago.

THE DATA REGULATION SNOWBALL
Respondents required to secure data within a regulatory framework in their organization

15% NO   8% DON'T KNOW

77% YES
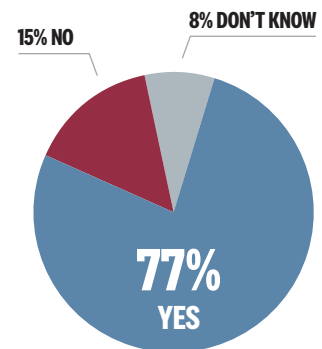
Respondents expecting to face new data privacy regulations in the near future in their organizations

20% NO   23% DON'T KNOW

57% YES
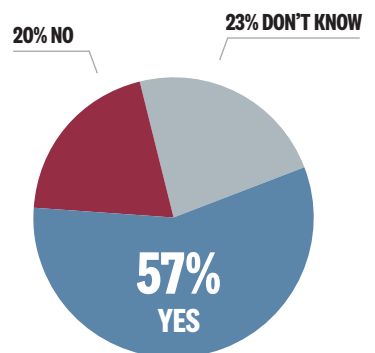
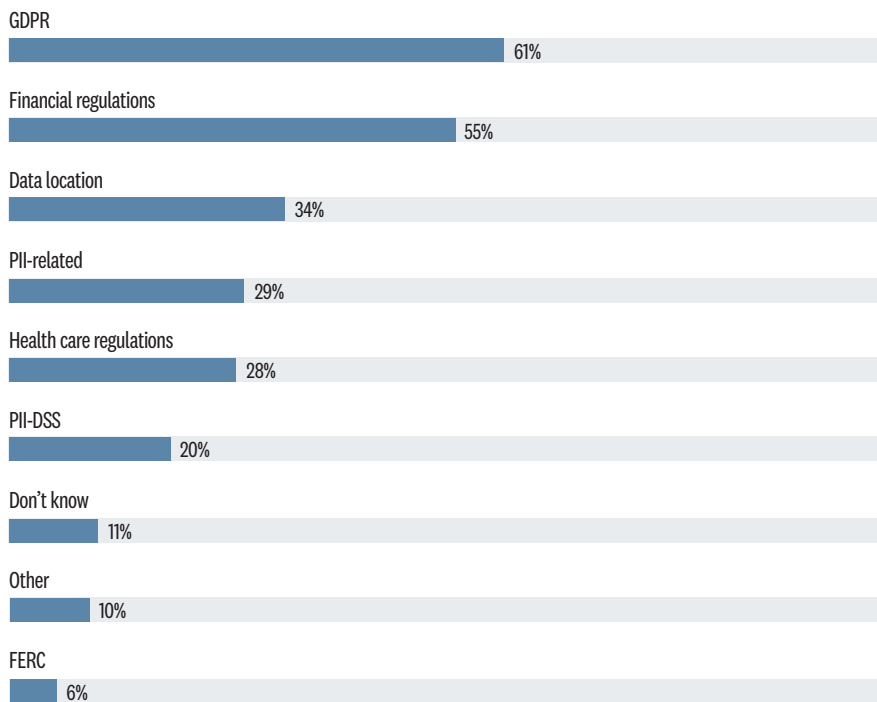SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

Claudia Imhoff, president of Intelligent Solutions Inc. strongly suggests creating an **enterprise data catalog**—not just for the data **but also for the algorithms** that are using it, the people who are using it, and the relevant privacy and security policies.

CAUGHT IN A GROWING REGULATORY WEB

Data sovereignty and/or data regulations affecting respondents' organizations

[SELECT ALL THAT APPLY]

GDPR
61%

Financial regulations
55%

Data location
34%

PII-related
29%

Health care regulations
28%

PII-DSS
20%

Don't know
11%

Other
10%

FERC
6%

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

agnosticism and a means of countering any fears of cloud vendor lock-in by offering a choice of CSP platform. Migrating from one DBMS to another is time-consuming and resource-intensive. If we can eliminate the effort to rewrite an application, and only focus on extracting data from one cloud to another, the task is radically simplified."[8]

The Opportunity Versus the Reality

To create intelligent products and services, improve automation, or achieve any future business objective that requires data, organizations will need to run a range of analytics with their data from wherever it is. Most are not ready.

Two-thirds of the executives surveyed say their organization needs a comprehensive data strategy in order to meet its strategic goals over the next three years, yet only one-third of them say their analytics and data management capabilities are on course to meet those goals.

Instead, Claudia Imhoff, president of Intelligent Solutions Inc. and founder of Boulder BI Brain Trust, sees "chaos" when she looks at the way many organizations are trying to manage their data, with some on-premises and some in multiple clouds or a hybrid environment. "When you scatter the data into various repositories, it can create a governance nightmare," she says. "We start to see data redundancy, we see errors, we see problems with even understanding where the data is, who's using it, how often they're using it." Some business units even end up paying for data already owned by their

Cloud Data Masters

What are the tangible benefits of an enterprise data strategy? A comparison of organizations that report a successful cloud data strategy shows that they are far ahead in embracing and embedding the kind of analytic tools and advanced capabilities necessary to become insight-driven organizations. These are the organizations that are not taking the risk of being left behind.

Only a third of respondents in the Harvard Business Review Analytic Services survey say they have an effective cloud data strategy, a contingent that qualifies as "cloud data masters."

Cloud data masters report that they...

- Are engaging in machine learning (47% vs. 25% for everyone else)
- Have developed AI capabilities (40% vs. 27%)
- Have met their goals for cloud migration (56% vs. 8%)
- Are moving data into the right environment (70% vs. 24%)
- Have developed capabilities to ingest data fast enough to make decisions in real time (43% vs. 13%)
- Have stream-processing capabilities (41% vs. 8%)
- Are on track to meet strategic goals for data capabilities (70% vs. 16%)

organization simply because they don't know where to find it.

Respondents acknowledge all of these issues—and more—as major challenges to gaining maximum value from data to meet strategic goals. More than half (54%) of executives surveyed report multiple data copies, 52% name organizational siloes, and 44% cite lack of interoperability. FIGURE 9

"Multi-cloud is not inherently bad. It's actually a nice way to store your data," says Imhoff. "But what I would advise anyone going down that path is you better have an architecture. You better understand what data resides where." She strongly suggests creating an enterprise data catalog—not just for the data but also for the algorithms that are using it, the people who are using it, and the relevant privacy and security policies. "The hope is that employees will use the catalog first to see if the data already exists somewhere before they pull out their credit card or begin to reinvent an existing database," she explains.
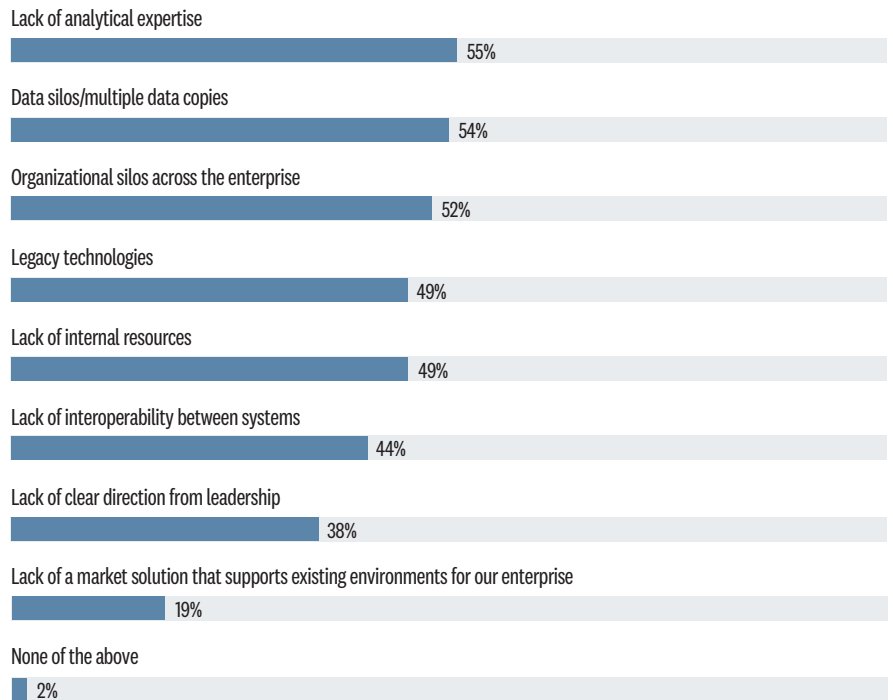
Managing data resources is a major part of any comprehensive data strategy, but it is only a part. Flynn at PwC points out the challenges of creating a data strategy for an organization that operates as a global network of firms. PwC has centralized capabilities and access management around some key services, but the firm must also maintain territory sovereignty in terms of where their data is located, the

BARRIERS TO MAXIMIZING DATA FOR STRATEGIC GAIN

Challenges your organization faces in gaining maximum value from its data to meet strategic goals

[SELECT ALL THAT APPLY]

Lack of analytical expertise	55%
Data silos/multiple data copies	54%
Organizational silos across the enterprise	52%
Legacy technologies	49%
Lack of internal resources	49%
Lack of interoperability between systems	44%
Lack of clear direction from leadership	38%
Lack of a market solution that supports existing environments for our enterprise	19%
None of the above	2%

SOURCE: HARVARD BUSINESS REVIEW ANALYTIC SERVICES SURVEY, JUNE 2019

TWO-THIRDS OF THE EXECUTIVES SURVEYED SAY THEIR ORGANIZATION NEEDS A COMPREHENSIVE DATA STRATEGY

way that applications can access the data, and the way that it's encrypted. Using a range of technologies allows PwC "to federate and deploy our cloud wherever we need it," Flynn says. "One tool can push containers of applications into a cloud environment, even if we can't actually see any of the data, or what they're actually doing in their environments," he explains. That particular tool allows the firm to manage applications and cloud usage centrally while leaving the data where it is.

The ultimate goal: to empower every team with the data and resources to develop the best intelligence for their business.

Conclusion

To gain competitive advantage from data assets requires an insight-centered strategy that embeds analytics capabilities across the organization. Acting on this strategy means developing the ability to apply a range of digital science tools—BI, machine learning, and AI analytics. It means being able to process and analyze data from multiple locations—on premises, across multiple clouds, or from a range of endpoints. It means establishing a way to apply a range of analytics on the same data set with a consistent security and governance policy. It means taking advantage of the best infrastructure options for agility and elasticity across multiple cloud environments. And, finally, it means maintaining an open platform so data is available where it's needed in a consistently secure and traceable way. The right strategy should make it easier for IT to control and easier for people to use—irrespective of where analytics run or data resides.

Endnotes

1   Gartner Press Release, "Gartner Identifies Top 10 Data and Analytics Technology Trends for 2019," February 18, 2019, https://www.gartner.com/en/newsroom/press-releases/2019-02-18-gartner-identifies-top-10-data-and-analytics-technolo.

2   Ibid.

3   Gartner, "Are You Ready for Multi-cloud and Intercloud Data Management?" Adam Ronthal, Rick Greenwald, Donald Feinberg. Published: 24 May 2019.

5. Ibid

6. Ibid

7. Ibid

8. Ibid

METHODOLOGY AND PARTICIPANT PROFILE

A total of 185 respondents drawn from the HBR audience of readers (magazine/
enewsletter readers, customers, HBR.org users) completed the survey.

SIZE OF ORGANIZATION

18%	**21%**	**10%**	**15%**	**7%**	**29%**
> 100 EMPLOYEES	100–499 EMPLOYEES	500–999 EMPLOYEES	1,000–4,999 EMPLOYEES	5,000–9,999 EMPLOYEES	10,000+ EMPLOYEES

SENIORITY

23%	**36%**	**24%**	**17%**
EXECUTIVE MANAGEMENT/ BOARD MEMBERS	SENIOR MANAGEMENT	MIDDLE MANAGEMENT	OTHER

KEY INDUSTRY SECTORS
OTHER INDUSTRIES WERE LESS THAN 7% OF THE TOTAL.

31%	**10%**	**9%**	**9%**	**8%**	**7%**
TECHNOLOGY	FINANCIAL SERVICES	BUSINESS/ PROFESSIONAL SERVICES	CONSULTING SERVICES	MANUFACTURING	GOVERNMENT/ NOT FOR PROFIT

JOB FUNCTION
OTHER INDUSTRIES WERE LESS THAN 6% OF THE TOTAL.

17%	**16%**	**9%**	**8%**	**7%**	**7%**	**7%**	**6%**
IT	GENERAL/ EXECUTIVE MANAGEMENT	STRATEGIC PLANNING	R&D/INNOVATION/ PRODUCT DEVELOPMENT	CONSULTING	ENGINEERING	SALES/BUSINESS DEVELOPMENT/ CUSTOMER SERVICE	INFORMATION/ KNOWLEDGE MANAGEMENT

REGIONS

38%	**27%**	**22%**	**8%**	**4%**	**1%**
NORTH AMERICA	ASIA/PACIFIC	EUROPE	LATIN AMERICA	MIDDLE EAST/ AFRICA	OTHER

Figures may not add up to 100% due to rounding.