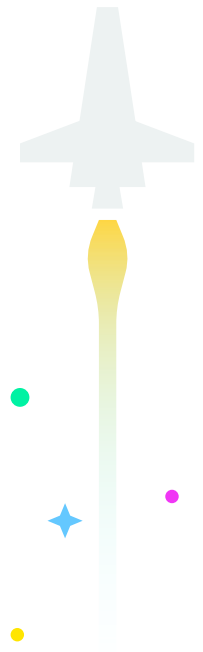


RESEARCH REPORT

The State of the Data Race 2022

Leveraging real-time data—for resilience in
the downturn and growth in the long term



SURVEY RESULTS FROM 500+ TECHNOLOGY
EXECUTIVES AND PRACTITIONERS ON HOW
REAL-TIME DATA DRIVES REVENUE GROWTH
AND INCREASES DEVELOPER PRODUCTIVITY

DataStax

Table of Contents

Executive summary	3
Introduction	6
Chapter 1: Real-time data is taking hold—and paying off	9
Chapter 2: Real-time data: a boon for developers	12
Chapter 3: There are hurdles – but leaders have cleared many of them	16
Chapter 4: It’s still early days for real-time apps	20
Final thoughts	22

Executive summary

The second annual State of the Data Race report is based on an in-depth survey of more than 500 technology leaders and practitioners across a variety of industries about their data strategies. This year's report focuses closely on how organizations use real-time data—the data that powers in-the-moment use cases such as recommendations and personalization, or always-up-to-date inventory and logistics—and found that leveraging it pays off in two important ways: innovations with real-time data lead to higher revenue growth and increased developer productivity.

Real-time data equals real success

There's a strong correlation between data leaders—organizations that excel at leveraging data to create new products and new revenue streams—and their use of real-time data. **Fifty-two percent of these organizations say they have a strategic focus that's built around using real-time data to create value across the organization.** There's a good reason for this: they are also most likely to say that real-time data has a transformative effect on revenue.

In other words, an organization-wide focus on leveraging real-time data is a proven way to accelerate revenue growth.



The increased likelihood of the “transformative impact” of real-time data on revenue growth from making it an organization-wide strategic focus

Succeeding with real-time data requires organization-wide commitment

The biggest hurdles organizations face when implementing real-time data are data complexity, costs, data accessibility, and finding the right talent for the IT team.

While today's data leaders have made countless investments required to overcome technical barriers, what do you think their number one challenge is? Thirty-five percent of leaders surveyed said that their top barrier to leveraging real-time data is finding the appropriate skills in their business units. These leaders are 1.5X more likely than the overall average to strongly agree that they have the right tools to build real-time applications but are constrained by the lack of in-house development resources to successfully implement real-time data projects.

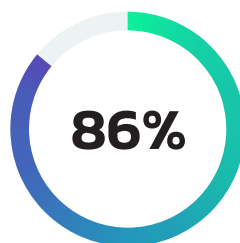
Therefore the first step to gaining competency with real-time data is to build the right technological foundation. Once it's properly addressed, then it's time to focus on enabling the teams most connected to customers—the business units—to build data products and drive revenue from real-time data.

Developers are getting real

Development teams are the key enablers to successful deployment of real-time data projects. They overwhelmingly rated interacting with real-time data and AI/ML, areas that are key to building in-the-moment applications that can respond intelligently and instantaneously, as leading areas of interest.

Eighty-six percent of developers at organizations with a strategic focus on real-time data say that “technology is more exciting than ever”—24 percentage points more than organizations with no real-time deployments.

Real-time data is critical to attracting talented developers, keeping them productive, and enabling them to make real business impact. Real-time applications, after all, are some of the most powerful ways to improve customer experience—and grow revenue.



The percentage of corporate developers that prioritize real-time data who say “technology is more exciting than ever.”

The race is on, and the prize is in sight

Many organizations are barely scratching the surface when it comes to harnessing the power of real-time data: the largest concentration of all organizations surveyed say they currently focus on internal uses (with the top example being monitoring and reporting on internal IT systems). But the potential of customer-facing deployments is crystal clear. While customer 360 and personalization initiatives have been around for some time, deploying them with real-time data has become a game-changer: organizations driving 20 percent or more of their revenue with data are 3X more likely to have deployed customer 360 or personalization than those driving 5% of revenue or less.

The bottom line: your customers will love real-time experiences. Your developers will love building them. And they can drive top-line growth. The time for real time is now.



From commitment to activation: Real-time data makes a real-time impact

Discovering what's about to become your new favorite album, from a band you've never heard of. A personalized offer and instant credit approval that makes giving the perfect gift at the right time easy. Immediate recommendations for in-stock alternatives to the items you were semi-panic buying.

In our daily lives, applications that use data to trigger helpful actions in real time can brighten our day or help us weather a storm. A real-world example: The Home Depot is using artificial intelligence to monitor weather activity and inventory, enabling the retailer to proactively move items like chainsaws or generators to where they'll be needed most.

More broadly, the transformation of businesses to drive actions with data in real time is [saving energy](#), [fighting fraud](#), and [making healthcare more efficient and affordable](#). And it's only just getting started.

The second annual State of the Data Race report serves as a survey-based validation for organizations and executives who've made a commitment to harnessing real-time data. In this report, you'll find that organizations that have prioritized real-time data as a strategic initiative report transformative revenue growth and increased developer productivity.

More than validation of digital leaders, however, this report gives any organization a playbook both to effectively use data to navigate what may be turbulent times ahead and to answer the question "what's next?"

For example, organizations expressing the most confidence in their ability to use data to increase revenue, reduce costs, and increase productivity are almost universally (90 percent or more) getting work done in cross-functional teams, aligned to business units.

Rightfully so, data is becoming a key tool in business leaders' toolboxes to help them hit their numbers in uncertain times. How do you squeeze cost and risk out of inventory and supply chain management? How do you gain a few points of uplift from shopping cart composition or your B2B sales operation? Many organizations find a role for apps that use data in real time to improve business processes or customer experiences.

With today's best-of-breed technologies for doing so available as a service, the timeline is months or quarters—not years. There are no capital costs and nothing to stop a proof-of-concept from getting started in minutes. And increasingly, developers are looking to work with exciting, cutting-edge technologies. Your developers are probably champing at the bit to do more with real-time data activation: in this survey, they rate working with real-time data and AI as among their most exciting opportunities.

If there's one line to take away in the current market context, it's "when the going gets tough, the tough get coding." The value of apps that use data to drive smart action in real time has been amply proven. The technology is ready, waiting, and openly available. Your organization's commitment, imagination, and drive are all that's needed to weather any storm on the horizon— and come out stronger for it.

A quick look back

The [first State of the Data Race report](#), published in 2021, explored the unique characteristics of "data leaders"—organizations that excel at using data to deliver value to customers and generate revenue. It was apparent that real-time data (or "fast data," as it was referred to in the first report) was table stakes for meeting market expectations. The goal was to validate patterns and practices beyond technology that set organizations apart as leaders, and uncover insights into steps that others less far along on their data-driven transformations could follow to accelerate their progress.

This new 2022 survey tracks several of the metrics that were researched in the first report and found that over the past year or so, organizations generally held the line with their use of real-time data.

For example, the percentage of respondents who say they could attribute more than 20% of revenue to their ability to capture and leverage data declined two percentage points (to 15%)—well within the margin of error.

One repeat question that did see a significant jump concerned the use of AI. Thirty-six percent of respondents say AI is in widespread production at their organizations. That's up from 28% in the first survey, and shows the progress that AI has made toward becoming a set of technologies that's understandable and accessible to a wider range of organizations.

About the research

In May 2022, **DataStax** and **Clearpath Strategies** surveyed 556 executives and technical practitioners at U.S.-based organizations. Most (84%) respondents were either key decision makers or had direct influence on the decision-making process; 29% were CIOs or CTOs, while nearly half held VP or director positions in a range of areas including operations, engineering, and cloud.

As with the previous report, respondents were segmented into five different categories (ranging from from “data leaders” down to “lagging”) based on the maturity of their organization in its data usage. The model used in this report is a refined version of the original one from the previous report. The original cluster analysis used a combination of 72 data-related behaviors, characteristics, technologies, cultures, people, processes, and business approaches.

The current model isolates the 12 most important variables from that original set and added a further two variables, which qualitative research suggested would have a high degree of correlation (quantitative analysis validated this hypothesis). A weighted scoring of respondents’ answers to these 14 variables determines how “mature” their organization’s data usage is and, therefore, which segment they fall in.

In this report, response intensity is typically expressed using the “top-box score” as a metric. Top box scores are often useful for detecting whether an attribute is a strong or differentiated characteristic of an organization or a standout priority or pain point for a respondent. A top-box score is usually represented by an answer of “strongly agree” to a particular statement or question.

The survey’s design and findings reflect both the quantitative research described above and on-going formal qualitative research and organic conversations. This includes dozens of interviews and conversations with executives (CIOs, CDOs, VPs of engineering, among others). Organizations in segment 1 (data leaders) or segment 2 were recruited for formal interviews.

Real-time data is taking hold—and paying off

“Real-time data is air. Real-time data is critical for everything, from your financials to customer sales to whether promos are working, whether advertising is working, whether the initiatives that you’re investing in are paying off. You’ve got to pivot so fast in today’s market. If you don’t have real-time data, you’re back in 1985.”

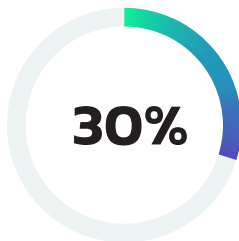
GREG SLY
SVP OF INFRASTRUCTURE
& PLATFORM SERVICES
VERIZON

Technology leaders and practitioners understand the importance of speed when it comes to staying competitive. Technology has evolved to the point where it can support that need for speed; companies of all sizes—not just the Netflices and Amazons—can build data architectures that can leverage data for powerful, in-the-moment user experiences.

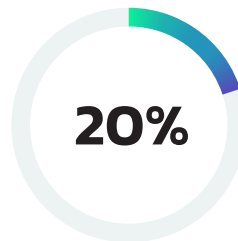
The survey data shows that for a significant percentage of organizations, the recognition of real-time data’s importance has converged with technological advances—and has become a strategic focus. Two big takeaways from the research: **a significant number of respondents have made real-time data a company-wide strategic focus, and this commitment contributes to an organization’s revenue growth.**

A sharp focus on real-time data

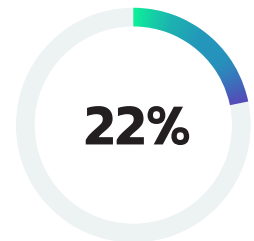
When asked to what extent their organization leverages real-time data, the largest percentage of those surveyed (30%) strongly agree that it is a strategic focus to use real-time data to create value across their organizations. Only 10% say their use of real-time data is minimal or nonexistent and that they didn’t plan to change that stance in the near future.



Our strategic focus is built around using real-time data to create organization-wide value



We use real-time data broadly in deployment, but it isn’t core to our strategy



We use real-time data in some relevant circumstances, but not all

When looking specifically at responses from leaders—organizations that excel at leveraging data to create new products and revenue streams—the strategic focus on real-time data is even stronger. More than half of these organizations (52%) say their corporate strategy zeroes in on building organization-wide value with real-time data.

Real time means real revenue

There's a strong correlation between building a real-time data strategy and revenue growth. Overall, 71% of all respondents agree (either completely or somewhat) that they can tie revenue growth directly to real-time data.

The effect on revenue growth is particularly pronounced at companies when they make real-time data a strategic focus. Respondents were asked which of several metrics (customer satisfaction, margins, customer loyalty, market share, and revenue growth) are impacted in a “transformative” way by an organization's real-time data capabilities. Among those organizations that make real-time data a strategic focus, 42% experience a transformative impact on revenue growth, compared to only 18% of those organizations that don't prioritize real-time data.

Why do applications powered by real-time data matter?

From a customer, employee, or partner perspective, benefiting from useful intelligence in real time is a key differentiator in fast-growing markets. In many industries, however, it's also becoming table stakes. Think of a bank with a credit card operation. The faster it can be alerted about suspicious activity in a customer account, the faster it can act to prevent serious damage—to a customer's balance and to the bank's reputation.

[Additional reading: “Turning the Bank Upside Down: How Data Speed Will Define Winning Financial Institutions” \(CIO.com\)](#)

78% of all respondents agreed that real-time data is a “must have,” not just a “nice-to-have”

Looking at this another way, organizations are 2.3X more likely to experience a “transformative impact” on revenue growth from real-time data when they make it an organization-wide strategic focus.

This gap between organizations that make real-time data a strategic focus and those who do not represents the potential revenue growth available to companies if they sharpen their attention on this important area.

Given the effect that it has upon revenue, it’s no surprise that 78% of all respondents agreed that real-time data is a “must-have,” not just a “nice-to have.”

How data leaders rank key metrics where their real-time data capabilities have had a “transformative” impact:



Hasn't “real-time data” been around for a long time? Why does it matter now?

The term “real time” has been around in one form or another for decades. “Reactive apps” were promoted as the next big thing in the early 2000s—but the technology wasn't really there to widely support the development of quick responses to user inputs on web or mobile.

Today, however, it's not a pipe dream. The technology (much of it in the form of open source software) is widely available. Organizations of all sizes can build applications that meet customer actions with a real-time response—so much so that users now expect that kind of instantaneous reaction from the applications they use everyday, no matter who's providing them.

Today, brand name disruptors like Starbucks, Netflix, Uber, Priceline, T-Mobile, and The Home Depot, along with smaller companies, like [Siggy.ai](#) and [Ryzeo](#), operationalize real-time data and build industry-disrupting businesses atop it. Those who aren't are feeling the pressure and scrambling to figure out how to compete with their data-first, digital rivals.

Real-time data: a boon for developers

Who does the most work with real-time data in your organization?

Developer teams

Data analysts

Data engineers

As noted in the previous chapter, the research reveals a strong correlation between leveraging real-time data and revenue growth—and the connection is particularly significant for companies that treat real-time data as a strategic imperative. There’s also another advantage that real-time data brings to organizations that make it a priority: it helps the people responsible for building instantaneous, in-the-moment experiences—developers—do their jobs better.

When asked how the use of real-time data has impacted developers’ jobs, an impressive 66% of real-time-data-focused organizations agree that developer productivity has improved.

This is particularly important because no one at the organizations surveyed works more closely with real-time data than developers. **The largest percentage of all organizations (27%) agreed that developer teams were the groups that worked most extensively with real-time data** (this percentage was just a bit higher—32%—for organizations that hold real-time data as core to their strategy).

Taken together, these two significant findings highlight the importance of ensuring that developers have fast and ready access to real-time data. For these builders, it’s all about ease of building and time to market. The complexity and siloed technologies that plague many organizations’ data estates can be a serious detriment; developers shouldn’t have to grapple with figuring out where the data they need resides, or how to access it.

For more on enabling developers with real-time data, read “You’ve Got the Data. Why Can’t Your Developers Build with It?” in CIO.com

But it's not just about enabling access to the data. It's also about re-orienting how organizations operate to make developers more productive.

Compared to organizations with less-aggressive commitments to real-time data, those who make creating value with it an organization-wide strategy are more likely to:

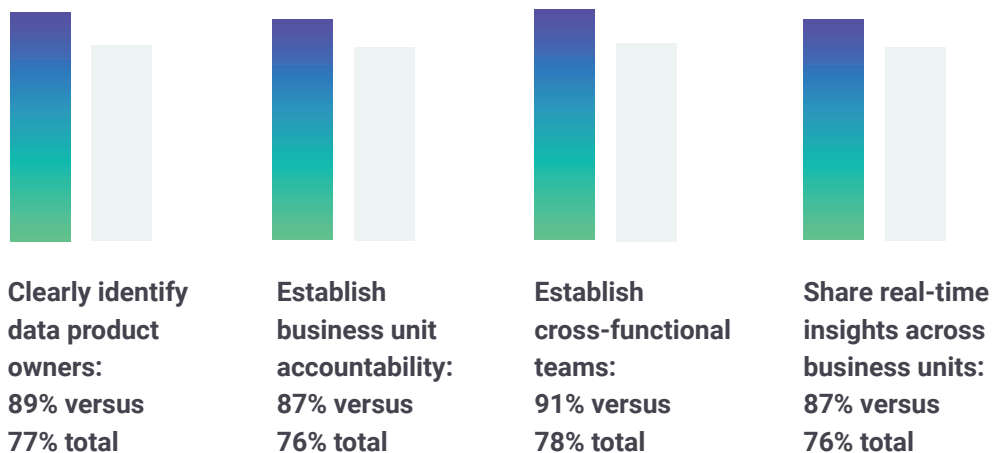
- Have clear product owners (43% versus 32%)
- Have business unit accountability of data (42% versus 30%)
- Have line-of-business staff, developers, and data scientists working together in cross- functional teams (45% versus 32%)

This organizational pattern lines up every role behind one North-Star goal on the way to creating value: *shipping an application*.

Until an application ships, all the good ideas in the world won't impact the customer experience—or the bottom line. Organizations that make a strategic commitment to creating value with real-time data are making the changes that focus attention, align incentives, and drive prioritization around the ground truth that **everyone will sink or swim together based on developers' productivity**.

Best practices for creating developer velocity building real-time apps

Organizations that have made creating value with real-time data a priority vs. all organizations surveyed



What kind of technology stack is necessary to take full advantage of real-time data?

There are three key elements required to deliver a real-time data experience. First, it's critical to have a data store that's optimized for customer context and instant access. It's how organizations take advantage of their real-time data "at rest." NoSQL databases are optimized for modern data applications that require large data volume, low latency, and flexible data models. [Apache Cassandra](#)® is an obvious choice, with its high throughput and ability to support applications that are globally distributed and always-on.

Secondly, streaming technology is required. Organizational behaviors and actions need to be visible and available to all applications across an organization; a best-of-breed streaming system should not only pass events along from one service to another, but store them in a way that keeps their value for future usage. The system needs to scale to manage tons of data and millions of messages per second—the kind of performance that real-time apps demand. [Apache Pulsar](#)™ is an advanced, open source streaming and messaging technology that's ideal for handling real-time data. It was built for the high throughput and scalability that many data-driven applications require.

Finally, it's critical to empower developers to make the most of real-time data—quickly and easily. An API layer lets them build applications with freedom of choice and without operational distractions. [Stargate](#), for example, is an open source data API layer that sits between applications and the database and offers a variety of endpoints for developers to build with.

[For more on Cassandra and real-time data, read "Two Reasons Why Cassandra is the Database for Real-Time Applications" in CIO.com](#)

Another big development benefit that accrues to organizations with a focus on real time: attracting and retaining strong talent. Five years ago, developers sought the freedom to create as they saw fit; when the [Stack Overflow 2017 Developer Survey](#) asked what they valued most in a new job, the largest percentage (53.3%) of developers said remote work options were a top priority.

With that freedom well-established (hastened in part by COVID-19), developers now want to work with cutting-edge technologies—and real-time data stands out as one area where the action is (35% of all developers polled described working with real-time data as “very important” – making it the second most important job feature after flexible and remote work opportunities).

The developers most tapped into next-generation technologies (those who describe themselves as the first in their organization to learn about new tools and technologies and those upon whom others rely on for answers about them) describe **interacting with real-time data and building AI and ML-powered apps as the most important factors in deciding where to work**. And 86% of developers at organizations with a strategic focus on real-time data say that “technology is more exciting than ever”—24 percentage points more than organizations with no real-time deployments.

Nothing kills developer enthusiasm like bureaucracy and work that doesn't add value. They love to ship code and build applications that have a real impact. Real-time applications and features are some of the most powerful ways to impact customer experience. So it's imperative for organizations to invest in making it easy for developers to access real-time data.

How a real-time data stack helped simplify development at Siggy.ai.

[Siggy.ai](#) is a real-time recommendation app that integrates with Shopify, the e-commerce platform for online stores and retail point-of-sale systems. Chang Xiao, Siggy.ai's founder and CEO, tried building a stack on his own using a combination of AWS and open source technologies. He began searching for an always-on solution that supported specialized operations while making development easy. Chang chose [DataStax Astra DB](#), and within weeks, he went from struggling with unreliable databases and server issues to building and delivering in-the-moment, AI-powered recommendations to shoppers everywhere.

There are hurdles – but leaders have cleared many of them

“A remarkably small number of companies have truly harnessed the power of all of their data to generate true revenue growth. Why? Companies’ most valuable data—generated through the billions of interactions they have with their customers—remains locked in silos and isolated in diverse operational data stores.”

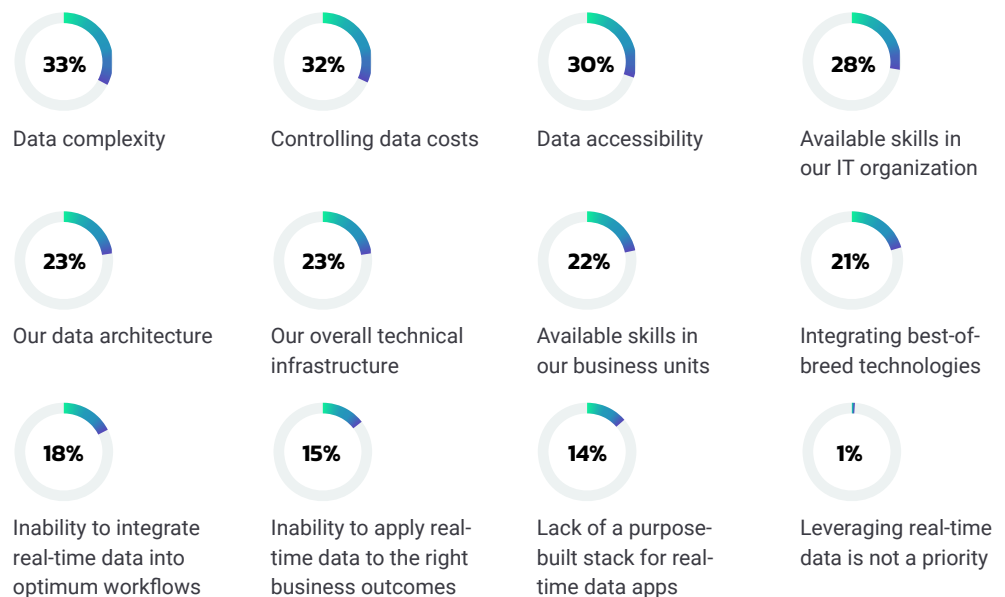
CODY SANFORD
FORMER CIO
T-MOBILE

Leveraging real-time data to take instantaneous action used to be a technology problem. The concept of real-time data has been around for a long time (some of you might recall the idea of “reactive apps” that surfaced in the early 2000s), but the technology wasn’t mature enough to enable the kind of in-the-moment experiences people enjoy when ordering a Starbucks drink or tracking the progress of an upcoming rideshare.

Today, the technology is widely available to organizations of all sizes. But implementing and benefiting from a real-time data strategy still carries challenges. The organizations surveyed said the biggest hurdles they face when trying to leverage real-time data are complexity, costs, and accessibility (39%, 32%, and 30%, respectively, identified these as top issues).

It’s no surprise. Over time many organizations have invested in a variety of point technology solutions. While this might have worked for one team or one project or one application, the end result of this effort was to lock data in a variety of silos across the organization.

The biggest barriers to leveraging real-time data



The top challenge for the most accomplished real-time data organizations is the availability of the necessary skills in their business units to leverage real-time data.

Cost is a significant challenge that can arise from data complexity. With a panoply of different technologies and data silos, enterprises have to maintain too many products and too many disparate skills; the costs of leveraging and scaling data can spiral out of control. Similarly, when real-time data is ferreted away in myriad repositories, without a holistic method for streaming it across an organization, developers find it exceedingly difficult to access the information they need to build.

[For more on data complexity, read "What Stands Between IT and Business Success?" in CIO.com.](#)

For leaders, the challenges lie with people, not tech

The challenges mentioned above don't rank as highly for data leaders—and the reason why has a lot to do with the progress these organizations have made in building real-time data architectures. **The number one issue for the most accomplished real-time data organizations is the availability of the necessary skills in their business units to leverage real-time data (the largest percentage – 35% – identify this as their top challenge).**

This makes sense. In the previous State of the Data Race report, 95% of data leaders say they've laid out crystal-clear ownership for turning data into products—a key philosophy when looking to leverage real-time data. Primary accountability for the data governance that creates the context for building products almost universally lies with business domain owners.

The challenge that leaders highlight in the latest poll is a strong indicator of the successful trail they've blazed with their real-time data architectures. They've built the technology foundation to successfully leverage real-time data, and now they're working on the next step. They understand that the right place to focus on building data products and the revenue they generate isn't in IT, it's within the business units.

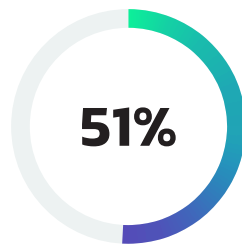
In a [recent article](#), Harvard Business School professor Marco Iansiti and Microsoft CEO Satya Nadella note that the companies experiencing the most success with data have a core of experts and broadly accessible, easy-to-use tools—but they also are investing heavily in building capabilities among large groups of people in lines of business. Organizations that follow this path can measure up to the likes of Uber or Airbnb when it comes to software-based innovation and companywide analytics capabilities, Iansiti and Nadella write.

When data leaders align technology infrastructure strategy to a business-domain-centric data strategy, IT's mission becomes the enablement of the highest-value uses of data identified by those with the deepest connections with customers.

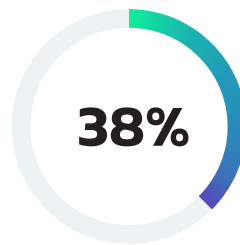
Interestingly, the biggest fear (by a significant amount) that organizations voice regarding the use of real-time data at their organizations are challenges around security. When one isolates this response for data leaders, the concern is even more significant: **51% of all organizations point to this as a significant fear, while the percentage is 61% for data leaders.**

There are two potential reasons for this. First, leaders have tackled many of the challenges faced by organizations that aren't as far along in their progress building their real-time data capabilities. Now they can train their focus on security. Secondly, real-time data and apps drive significant customer satisfaction and revenue—to the point where these metrics are business-critical. **The stakes are simply higher for data leaders; any security issues that might crop up with real-time data could threaten a core part of the business, not just a side project.**

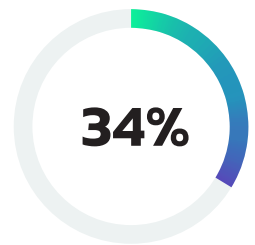
What's the biggest fear regarding the use of real-time data at your organization?



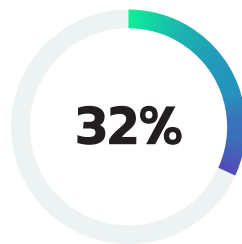
Security challenges



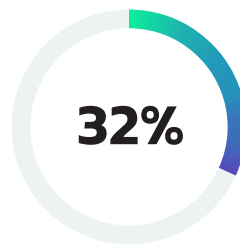
Too much complexity



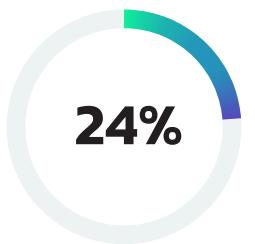
Compliance issues



Architecture limitations



Lack of skills to execute



Failure to deliver competitive customer experience

The bottom line: There's a clear playbook for success

There are challenges when it comes to deriving value from real-time data, and those challenges evolve as organizations progress toward real-time data proficiency. The good news: there's a clear playbook for surmounting the hurdles on the path to success. The challenges are well-understood and have proven to be solvable.

First, there's assembling the technology stack to power real-time data applications. Leaders have laid this groundwork—and they're attacking the next set of challenges with gusto.

Second, there's the imperative for everyone—not just developers and engineers, but also business leaders—to use real-time data to drive the business and grow revenue.

Consider pharma giant Novartis (as detailed in the Harvard Business Review article mentioned above). Over the past decade, the company invested heavily in data platforms and data integration. But it found that these investments only resulted in spotty success. Data scientists had little visibility into the business units, and, conversely, leaders from sales, supply chain, HR, finance, and marketing weren't embracing the available data.

Once data scientists were paired with business employees with insight into where efficiency and performance improvements were needed, and once frontline organization employees were trained to use data for innovation, the intensity and impact of transformation accelerated.

It's still early days for real-time apps

As mentioned earlier, the technology is readily available now to help organizations take advantage of real-time data. From high-throughput, scalable datastores, to streaming technologies, to simplified ways for developers to interact with real-time data—the building blocks, including many open source options, are here now.

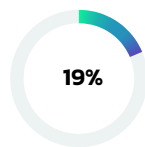
Organizations of all sizes can build applications that meet customer actions with a real-time response—but, according to the survey findings, that's not yet where the focus has been. Rather, most companies have been going after low-hanging fruit.

The two top use cases that organizations use real-time data for are internal: monitoring and reporting on internal IT systems (46%) and monitoring user/customer activity (43%). These areas are often associated with real-time data scored significantly lower. Thirty-two percent surveyed said they use real-time for improving customer personalization and customer experience; customer 360 solutions was chosen by 19%.

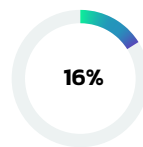
Among the data leader companies, the top two ways their developers interact with real-time data are in-the-moment analysis and building real-time features and apps (72% and 63%, respectively).

The top five real-time data use cases for data leaders

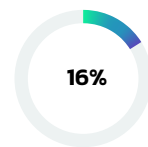
Compared to the average respondent, data leaders are:



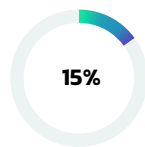
More likely to use real-time data to monitor internal IT



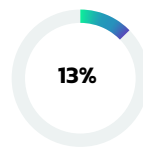
More likely to use real-time data for customer personalization



More likely to use real-time data for location/logistics



More likely to use real-time data for chatbots/messaging



More likely to use real-time data for payments/processing

This is consistent with the best practices for developer velocity highlighted in chapter two. The first step is aligning all roles—not just developers—to focus on how activating data in real time can improve business processes or customer experiences. The second—and necessary—step is laying the groundwork for real-time analysis and hypothesis testing. This unlocks the opportunity to take human intervention out of the equation and close the gap between an event occurring and an impactful action. This should be a consistent, repeatable pattern across an organization.

The future looks bright

Although there's a variety of challenges that organizations face when it comes to implementing real-time data—including complexity, accessibility, and cost—none of these seem insurmountable to organizations that have begun to embrace real-time data. They are quite optimistic about expanding its role to support continued business growth. Looking ahead, a majority of leading organizations (61%) expect “much more” usage of real-time applications over the next three years.

Why open source?

You probably noticed that three key pieces of the real-time data stack highlighted earlier are open source. The most popular real-time data technologies grew out of the “best of the Internet” to solve novel problems about scale and data velocity. Cassandra was developed at Facebook to manage the reams of data from the messages users exchange over the social network. Apache Kafka was built at LinkedIn to handle real-time data feeds from activity tracking and application metrics. Each of these technologies were open sourced once their designers understood the potential for them to benefit a wider range of applications.

The open source cycle of innovation continues as the importance of real-time data keeps growing. Pulsar, for example, emerged in recent years as a more cloud-native friendly alternative to Kafka (it was open sourced by Yahoo! in 2016).

Joining the open source ecosystem means avoiding reinventing the wheel. Why does this matter? Because what sets an organization's real-time data experiences apart won't be the infrastructure. It's people, applying their domain knowledge to identify opportunities to build real-time, high-growth applications and create new ways to delight users.

[Additional reading: “How OSS Fuels Enterprise Innovation” \(CIO.com\)](#)

Final thoughts

Applications that learn minute-by-minute and can trigger smart actions hundreds of times per second or millions of times a day have proven their value for commerce, entertainment, and business operations. And there are good reasons to believe the best is yet to come.

As this report shows, some powerful, complementary forces have converged. The open source cycle of innovation is as strong as ever—and now the best-of-breed tools it has generated are available on demand, as-a-service, to everyone.

Cloud tools and DevOps processes have made developers more productive—and now leading organizations have shown that embedding developers in cross-functional teams that are empowered to drive an organization forward is a best practice for accelerating the activation of data in real-time. Organizations that commit to leveraging real-time data are shipping the winning applications that it powers—and seeing transformational results.

This is a time of big challenges and elevated uncertainty. But there are two things that look to be smart bets today.

One: you, your partners, and your stakeholders (and for that matter, your competitors) will have more data at your disposal than ever before, by a longshot.

And two: those organizations that commit to asking the right questions, picking up the right tools, and empowering their people to use them can and will find ways in which putting it to work in real time will drive revenue and growth.

ABOUT DATASTAX

DataStax is the real-time data company. With DataStax, any enterprise can mobilize real-time data and quickly build the smart, highly scalable applications required to be a data-driven business.

The DataStax Astra cloud service uniquely combines the power of Apache Cassandra®—the world’s most scalable database—with the advanced Apache Pulsar™ streaming technology in a unified stack, available on any cloud. Hundreds of the world’s leading enterprises, including Verizon, Audi, ESL Gaming, and many more rely on DataStax to unleash the power of real-time data to create the in-the-moment digital experiences that can win new markets and change industries. Learn more at DataStax.com.

© 2022 DataStax, All Rights Reserved.

DataStax, is a registered trademark of DataStax, Inc.. Apache, Apache Cassandra, and Apache Pulsar are either registered trademarks or trademarks of the Apache Software Foundation.