

Ultra-Fast Data Access Is The Key To Unleashing Full Big Data Potential

Executive Summary

The proliferation of big data generated by enterprise applications, consumer web/mobile apps, and the Internet of Things (IoT) has afforded an unprecedented opportunity for businesses to know more about their customers than ever before. Interest and excitement about big data is on the rise, and online public discussion about big data has steadily increased over the past year as the market has become more passionate about big data and the prospect of using it.¹ However, the unfortunate truth is that most of the potential of big data lies dormant, as most businesses lack the tools and capabilities to effectively access, process, and analyze the data available to them in a timely manner. In an age where a deeper understanding of customers is a competitive necessity, this inability to effectively analyze all data represents a significant missed opportunity and, perhaps worse, a competitive disadvantage for firms whose competitors are analyzing big data.

Analyzing big data is critical for enterprises to serve customers and improve overall business performance. Historically, data access, processing, and analysis were most important for reporting on business performance, followed by customer insights and business intelligence, but our study revealed that customer insights will replace business performance as the top priority for data processing and analysis. This supports Forrester's overarching assertion that enterprises have entered the age of the customer — an era in which customer obsession is critical to success and businesses must focus more effort on winning, serving, and retaining customers.

In September 2015, SAP and Lenovo commissioned Forrester Consulting to explore the hypothesis that most businesses are only analyzing a small part of the data available to them, the result of which is significant missed opportunities to better serve customers and improve business outcomes. To test this hypothesis, Forrester conducted an online survey of 100 IT decision-makers in enterprise organizations in the US. Respondents were responsible for decisions related to data processing and analytics and were at companies that were currently using Hadoop to manage their data. As well, Forrester utilized big data in this research by leveraging NetBase to scrape all social media content related to big data (including “big data” and #bigdata) that IT professionals and consumers posted publicly on the Web between December 2014 and December 2015. Content included anything posted across forums, blogs, microblogs, comments, consumer reviews,

and social networking sites to help assess relevant conversation about big data.

KEY FINDINGS

Forrester's study yielded four key findings:

- › **Customer insight is the top priority for data processing and analysis.**
- › **Seventy percent of available data is unanalyzed.**
- › **Enterprises are adopting Hadoop but continue to face persistent challenges with data access, preparation, and analysis.**
- › **Businesses want faster processes and time-to-insight as they look to improve big data practices.**

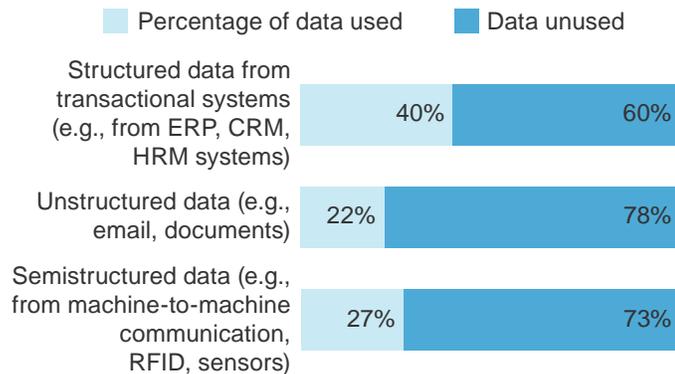
Current Big Data Capabilities Are Sorely Lacking

The digital age continues to be the catalyst for expanding data volumes across all enterprises. Our custom survey of enterprises that are using Hadoop to store and process big data found that 83% of companies are storing between 100 to over 500 terabytes of data. Further, the survey showed that the source of the data is actually a near-even split between internal data (e.g., ERP and business applications) and external data sources (e.g., social and connected device data).

That's plenty of data, but here's the rub — most of that data is unused for analytics. Forrester data shows that an average of 70% of data within the enterprise goes unused (see Figure 1).

FIGURE 1
Most Big Data Goes Unused

“Please estimate what percentage of the total size/ volume of data within your company is currently using for BI.”
(Mean percentage)



Base: 639 IT and business decision-makers at enterprise companies
Source: Global Business Technographics® Data And Analytics Survey, 2015, Forrester Research, Inc.

DATA ACCESS IS CRITICAL

For most businesses, the full value of using data comes from accessing and analyzing the data from numerous existing transactional systems together with new sources of external data to paint a more complete and valuable picture of customers and the business. Analyzing the data sources

separately provides siloed and incomplete insights, thus driving the need for improved data unification and access capabilities. Our custom survey found that for 80% of companies surveyed, data requests often utilize multiple data sources. But only 21% of companies strongly believe they are using their internal data to its fullest potential. Those that believe they are fully utilizing external data is even less, at 14%.

“The biggest problem with big data is using it.”

— User-generated social media post, based on Forrester’s analysis of NetBase aggregated social listening data, Dec 2014 to December 2015 (global)

The First-Order Big Data Challenge Is Ubiquitous Access To Data

Using big data exacerbates the challenges of data access, processing, and analytics. A recent Forrester report postulated that “a good rule of thumb in any BI (and actually most ERP, CRM, and other enterprise business software) initiative is to assume that 80% of the effort is going to center around data integration activities.”² It is no surprise to see that the top challenges reported with using more data relate to data access and preparation issues. The top challenges were:

- › **Difficulty accessing data from multiple sources.** A typical enterprise has a large portfolio of dozens or even hundreds of internal business applications that contain valuable data that is difficult to access in a unified way for analytics. Analyzing that internal data together with the growing number of external data sources such as IoT is essential to achieve the deepest possible insights. A conversation from an online forum highlighted this challenge, as a user posted the following: “I encounter a big problem with big data in my SQL. We have two big tables with the same structure. . . . We want to union the two tables to get a clean version by removing data redundancy. However, we find this task not easy with the challenge of big data.”

- › **Too many data formats to unify effectively.** In addition to enterprises having difficulty accessing a high volume of data from various sources, they also find it challenging to unify data with different formats, as many businesses lack the proper tools to do so effectively. Disparate formats from these different sources are typically housed in different databases and must be normalized to a common format before analytical data sets can be prepared. This challenge will be exacerbated as the number of data formats continues to increase with the expansion of IoT and as more and more connected devices come to market, all collecting data in varying formats. In fact, our social listening study showed that IoT is among the top related terms and hashtags used in big data discussions online, which validates the need for effective methods to manage various data formats.
- › **Difficulty in preparing data for analysis.** Most business users lack the proper skills to identify and discover the data they need amid the large volume of available data. As such, data requests are frequently funneled through IT staff, who then compile the data from necessary systems for the business users. Forrester data shows that the challenge with this process is that over 50% of companies

report it taking months or longer for IT to turn around big data analysis requests for provisioning new data sets for self-service analytics. The inability to quickly provision needed data impedes effective and timely use of data for analysis.

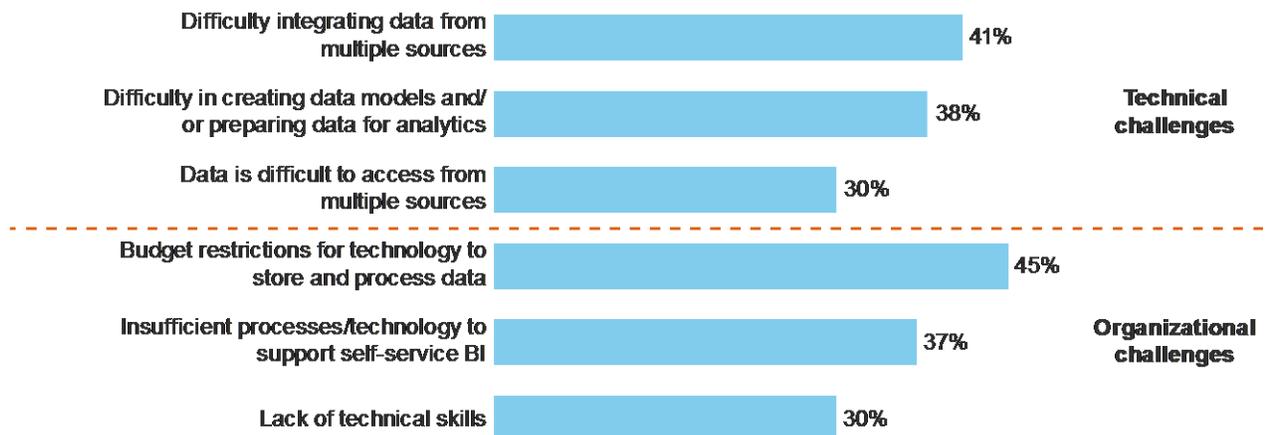
While data access challenges hinder more complete data usage, on the organizational side the lack of greater data usage is primarily the result of having insufficient tools and processes in place (see Figure 2). The top influencing factors included:

- › **Budget restrictions on new technology for data storage.** Budgetary issues were the top organizational challenge, with 45% of survey respondents indicating this was an issue. Without budget to purchase the proper tools and database systems to manage the growing volumes and varieties of data, enterprises will be limited in how effectively they can utilize their data.

FIGURE 2

Businesses Face A Number of Technical And Organizational Challenges Harnessing Big Data

“What are the challenges impeding you from processing and analyzing more data?”



Base: 100 IT and business decision-makers at enterprise companies
(not all responses shown)

Source: A commissioned study conducted by Forrester Consulting on behalf of SAP, September 2015

- › **Insufficient processes/technology to support self-service.** Enabling self-service data processing and analysis requires clearly defined processes and access methods such as SQL that can easily be leveraged without advanced training. The challenges of self-service are compounded by the data access challenges highlighted earlier, which all but ensure a dependence on IT for discovering and provisioning data unless new technology or processes are introduced.
- › **Lack of technical skills.** Having the right skills can pertain to both the IT staff charged with managing data and the business users who are looking to identify and explore data on their own. While enabling more self-service is the idea, business users must be able to access the data using tools they are already familiar with.

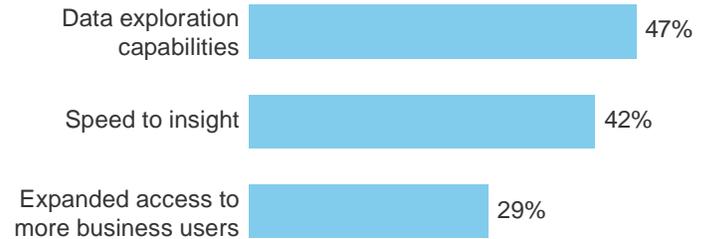
Ubiquitous, Fast Access To Data Is Key

As enterprises work to achieve the results of gaining deeper customer insight from big data, the first step is to ensure that all the data is available, prepared, and easily accessible. Customer preferences are individual and constantly in flux. Many customer insights are also perishable. They need to be discovered in a timely manner or risk becoming irrelevant. The weeks or months required for IT staff to provision data for analysis are sorely insufficient to capture valuable customer insight that is needed immediately. Our custom survey responses confirmed the need for this shift — when asked to consider what improvements would be most important for enhancing data processing and analysis capabilities, respondents cited these three priorities (see Figure 3):

- › **Real-time, comprehensive data exploration capabilities.** As enterprises experience expanding data volumes and more challenges accessing disparate data structures, they need to improve their ability to identify and discover data specific to their needs quickly and effectively.

FIGURE 3
Enabling Faster Access To Data Is A Key Capability Moving Forward

“As you look to improve your data processing and analytics capabilities, what capabilities are most important to your business?”



Base: 100 IT and business decision-makers at enterprise companies
Source: Global Business Technographics Data And Analytics Survey, 2015, Forrester Research, Inc.

- › **Faster insights, in time for actions.** Data democratization (i.e., empowering business users with the tools they need to discover their own data for analysis) enables businesses to facilitate more timely use of data, thus improving business agility by allowing them to take actions on those insights.
- › **Democratized access to more business users.** With improved exploration capabilities and familiar tools, enterprises can enable business users to explore data on their own and bypass the bottlenecks frequently caused by IT for data provisioning requests.

Key Recommendations

Big data is only going to get bigger and richer, as well as originate and flow from an increasing number of sources, both internal and external. Enterprises need a clear big data strategy that includes ultra-fast, ubiquitous access to all data in order to perform all types of analytics. The painfully slow, legacy-style data access and processing strategies of the past are insufficient. A modern data analytics strategy must provide a ubiquitous, real-time data access layer to all relevant data from all different sources. Firms must continue to leverage open source technology such as Hadoop and Spark for batch processing. But they also must create a “speed layer” that leverages in-memory technology to provide seamless, real-time access to all sources of data, whether or not they are in a Hadoop data lake. An in-memory platform can remove data latency by providing the power and speed to analyze more types of data in a single unified environment, thereby enabling faster processes. This ultra-fast speed layer powered by in-memory technology will allow firms to capture and act on valuable perishable insights across the entire business.

Appendix A: Endnotes

¹ Source: Forrester’s analysis of NetBase aggregated social listening data, December 2014 to December 2015 (global).

² Source: “Boost Your Business Insights By Converging Big Data And BI,” Forrester Research, Inc., March 25, 2015.

ABOUT FORRESTER CONSULTING

Forrester Consulting provides independent and objective research-based consulting to help leaders succeed in their organizations. Ranging in scope from a short strategy session to custom projects, Forrester’s Consulting services connect you directly with research analysts who apply expert insight to your specific business challenges. For more information, visit forrester.com/consulting.

© 2016, Forrester Research, Inc. All rights reserved. Unauthorized reproduction is strictly prohibited. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change. Forrester®, Technographics®, Forrester Wave, RoleView, TechRadar, and Total Economic Impact are trademarks of Forrester Research, Inc. All other trademarks are the property of their respective companies. For additional information, go to www.forrester.com. [1-WL7CPU][1-V7P0OR]
