



IT INDUSTRY TRENDS



BIG DATA...THE NEXT FRONTIER

Space: the final frontier. To paraphrase Star Trek, Big Data may help your business boldly go where no business has gone before. And while it may not be the final frontier, it is certainly the next frontier in IT.

Much like space, Big Data can be conceptualized in three dimensions (known as the “3 ‘V’s”):

- Volume: huge data sets
- Velocity: speed of real-time data in and out
- Variety: new sources and ranges of data types

Organizations that can properly harness these components of Big Data will gain a significant competitive advantage in the marketplace. So, like the starship Enterprise, this paper will explore huge data sets, seek out new sources and new ranges of data, and travel at warp speeds.

In August 2013, TEKsystems surveyed more than 2,000 IT professionals and more than 1,500 IT leaders on the topic of Big Data. IT professionals provided specific insight from the employee’s perspective on Big Data projects, while IT leaders provided insight into how Big Data will drive business growth. We asked each group to share their viewpoints on the Big Data landscape, the business impact of data and analytics, and the challenges most organizations face when implementing Big Data initiatives.

The Big Bang: Exploring the Big Data Universe

We live in an age where technology allows data to be collected more easily and at faster speeds than ever. About 90 percent of the world's data has been generated over the last two years. By the time you finish reading this article, people will have initiated 20 million Google searches, sent 2 billion emails and posted more than one million tweets. That's a mind-boggling quantity of data. But how "big" is Big Data, in terms of dollars? The research firm Gartner estimates that the direct and indirect worth of Big Data will reach more than \$230 billion by 2016. While maybe not as big as the final frontier, Big Data is truly a colossal market.

"Big Data" refers to huge data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a data set needs to be—i.e., we don't define Big Data in terms of being larger than a certain number of terabytes. We assume that, as technology advances, the size of datasets that qualify as Big Data will also increase.

While the quantity of data being collected increases every day, the ability to mine this information for insights is diminishing. Part of the problem occurs because Big Data sets come in so many different forms that it is impossible to process the information in a traditional relational database. The sheer volume, variety and velocity of the data, along with inadequate planning and evaluation, data management and skill shortages, combine to create a Big Data problem.

Traveling through Three Dimensions: Volume, Velocity and Variety

Survey respondents rank volume as the area of Big Data that causes the most difficulty. While the mountains of available data beckon business leaders to take advantage of it, that same volume presents the first obstacle in implementing Big Data initiatives. In order to take advantage of massive data sets, an organization first needs to determine how to store all of the information.

But even once that problem is solved, how does one start to process mind-numbing volumes of data? Fortunately, organizations have options available, ranging from storing data in the cloud to leveraging a Hadoop platform.

Once the volume problem has been addressed, organizations will need to tackle the other dimensions of Big Data. Factors such as variety in the types of data and the velocity with which it arrives also determine how organizations build and support IT infrastructure to support Big Data initiatives. To address these issues, IT providers will need to work closely with the business leaders to fully understand their goals and objectives and appropriately capture the necessary technical requirements. The data's complexity and heterogeneity, as well as the need for speed, will create immense pressures on IT infrastructure. If organizations want to capture the full value from Big Data, they will need to install new applications such as storage, computing and analytical software. They will also need to implement processes and techniques to analyze the data in a way that best serves the organization.

Much like space, Big Data can be conceptualized in three dimensions:



46%

Volume:
huge data sets



41%

Velocity:
speed of real-time data in and out



27%

Variety:
new sources and ranges of data types

Resistance is Futile: Big Data Is Upon Us

Although it is one of the hottest topics in the IT industry, Big Data is not a fleeting trend. Data and information are a “new class of economic asset, like currency or gold.”¹ Mastering this data frontier can provide businesses with insights and predictive analytics that will create a competitive advantage and could drastically alter the business landscape. IT leaders believe investments into Big Data initiatives will yield excellent return on investment, but reaching a state of Big Data nirvana won’t be easy. Leaders need to be comfortable in the new paradigm of “unknown unknowns,” creating tools for finding answers to questions that have yet to be asked.

The areas that will most challenge organizations include:

- Using new applications, processes and techniques: Organizations will need to evaluate a myriad of software and technology platforms, then install processes and techniques for optimizing the new technologies.
- Collaborating across departments: Big Data projects will require teams from different business units to work together, bringing into focus the entire data picture in ways that will offer insights previously unattainable.
- Reporting and recording of information (metadata): Maintaining accurate records about the information and data an organization captures or produces is critical, because poorly managed metadata can have significant negative effects on final reports.

Manage Data and Prosper: Data Management Policies

As organizations continue to generate data, the data itself becomes a challenge. Big Data projects require data to be accessed and shared across departmental functions. Organizations need to consider privacy, security, intellectual property, and ownership and liability factors when updating or implementing Big Data initiatives.

Ownership, within the organization, of the data and the possible impact on data integrity present another challenge. About half of our survey respondents believe the quality of their organization’s data is questionable. Identifying ownership and holding owners accountable to quality and security standards must be a fundamental aspect of any data policy. But doing so may be easier said than done. More than half of IT leaders (57 percent) and IT professionals (52 percent) report they don’t always know who owns the data. If one doesn’t know who owns the data, there is no one to hold accountable for its quality. As different sources and varieties of data are fused together for Big Data projects, ensuring the accuracy and quality of the data will be critical to their success

57%
of IT leaders



AND

52%
of IT professionals



report they don’t always know who owns the data.

Implementing a clear, well-communicated data policy will ensure Big Data initiatives address privacy, security, intellectual property, ownership and liability concerns. The policy will also safeguard quality so projects are not standing on a shaky foundation of faulty data. Then the organization must identify a team or specific individual such as a CDO (Chief Data Officer) who will identify inhibitors of Big Data initiatives, assign ownership of data sets and define the organizations’ strategy. The team or individual can then make recommendations on how the organization should address the challenges inherent in Big Data implementations.

1. World Economic Forum, 2013

Hailing on All Frequencies: Shortage of Skilled Resources

Defining challenges, evaluating software tools and implementing policies are just the first steps of the Big Data voyage. People are the key to realizing the full potential out of data initiatives. Organizations need the best people with the right skill sets to successfully realize any business objective, but securing top talent is always a challenge. Big Data projects magnify that challenge because of the broad range of skill sets needed. Eighty-one percent of IT leaders and 77 percent of IT professionals believe there is a significant shortage of workers with the skill required to plan, execute and take advantage of the potential of their organization's data assets. Some skill sets and roles, like data scientists or graphic designers, don't fall neatly into the world of IT, while others, like data forensic librarians or chief data officers, did not exist until recently. Tackling Big Data projects might mean rethinking job roles and titles, as well as the non-technical skills needed to make the best use of the data.

Key skills needed to take full advantage of Big Data include:

- Strong aptitude for business, technology, mathematics and statistics
- Knowledge of data visualization such as heat maps to analyze and present complex trends
- Hadoop: an open-source Apache framework to analyze and mine big data sets
- Programming skills such as SQL, Python, Unix, PHP, R and Java
- Teamwork: Big Data projects require collaboration between managers, IT administrators, programmers, statisticians, graphic designers and experts in the company's products or services
- Experience with data mining, modeling and hypothesis generation in support of high-level business goals

Finding the right talent is just one piece of the Big Data puzzle. Organizations need to consider whether they will leverage permanent, contingent, project-based or

fully outsourced models to address their needs; the best option is most likely a blended approach. Organizations will likely need a combination of well-trained staff, permanent and contingent labor to address various elements as the organization plans and builds its Big Data initiatives, and then integrates them into normal business operations. Some roles, such as a data scientist, would better serve the organization in a full-time capacity. But it might make more sense to hire positions like developers on a contingent basis until their portion of the project is complete. Organizations will need a mix of well-trained staff, permanent and contingent labor to address various elements as the organization plans, builds and runs its Big Data projects. Partnering with a vendor experienced in strategic workforce planning will ease the burden on internal staff and set the organization on the right path to Big Data success.

Organizations lack visibility into their end-user activity.



81% of IT leaders



77% of IT professionals

claim their site visits are not monitored.



The Next Frontier: Conclusion

As an organization embarks on a Big Data mission, leaders would do well to take on the mindset of a space traveler exploring the nearly infinite possibilities of the unknown. The ability to leverage huge volumes of data at incredible velocity, paired with the ability to continually evolve with business needs, will yield competitive advantage. But organizations must properly plan for success, get ahead of data quality and ownership challenges, and proactively address their talent sourcing and management strategies.

Armed with the extraordinary insights that Big Data can provide, leaders can innovate and make forward-thinking decisions based on real-world data.