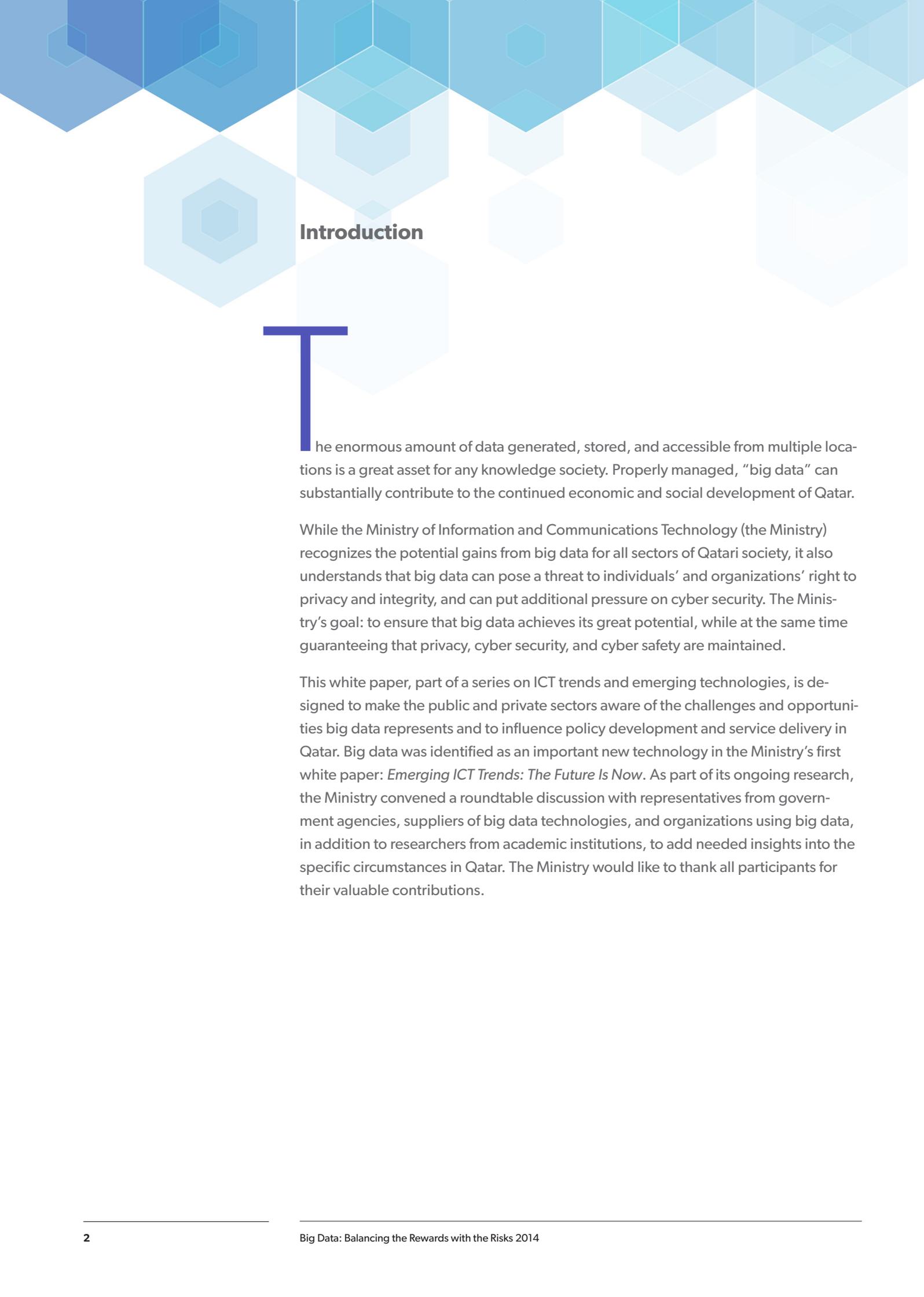


# BIG DATA: BALANCING THE REWARDS WITH THE RISKS

2014



## Introduction

The enormous amount of data generated, stored, and accessible from multiple locations is a great asset for any knowledge society. Properly managed, “big data” can substantially contribute to the continued economic and social development of Qatar.

While the Ministry of Information and Communications Technology (the Ministry) recognizes the potential gains from big data for all sectors of Qatari society, it also understands that big data can pose a threat to individuals’ and organizations’ right to privacy and integrity, and can put additional pressure on cyber security. The Ministry’s goal: to ensure that big data achieves its great potential, while at the same time guaranteeing that privacy, cyber security, and cyber safety are maintained.

This white paper, part of a series on ICT trends and emerging technologies, is designed to make the public and private sectors aware of the challenges and opportunities big data represents and to influence policy development and service delivery in Qatar. Big data was identified as an important new technology in the Ministry’s first white paper: *Emerging ICT Trends: The Future Is Now*. As part of its ongoing research, the Ministry convened a roundtable discussion with representatives from government agencies, suppliers of big data technologies, and organizations using big data, in addition to researchers from academic institutions, to add needed insights into the specific circumstances in Qatar. The Ministry would like to thank all participants for their valuable contributions.



## Background

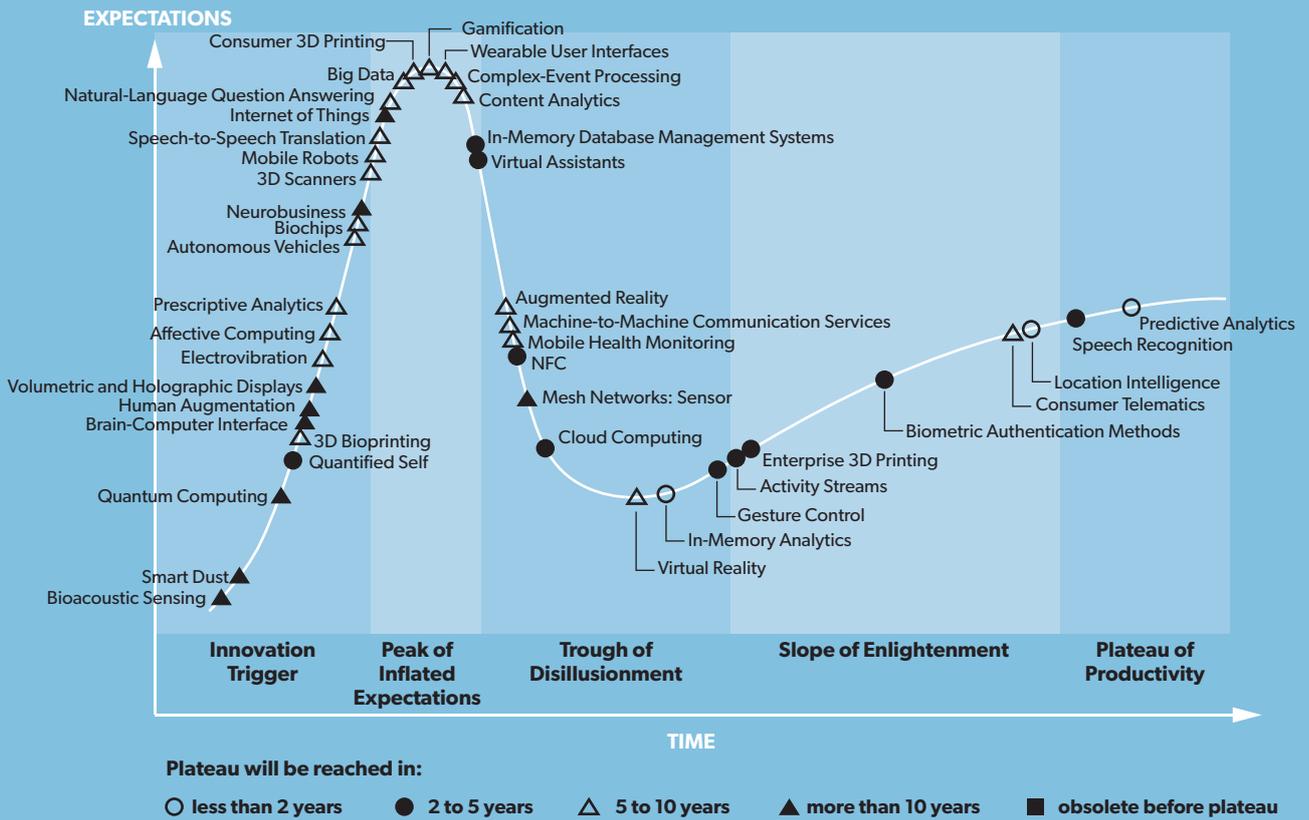
**By 2018 each smartphone is expected to generate 2 gigabytes of data every month.**

Organizations, individuals, machines, and devices are generating enormous amounts of data at an accelerating pace. On a global level, the National Institute of Standards and Technology—a US agency that works with industry to develop and apply technology, measurements, and standards—estimates that there will be some 7 zettabytes of data generated in 2014 (with 1 zettabyte equaling 1 billion gigabytes). Furthermore, by 2018 each smartphone is expected to generate 2 gigabytes of data every month. The rapid growth of data generation is attributable to the proliferation of Internet-connected devices and systems, the switch from analog to digital technologies, and the rapidly increasing use of digital media by organizations and individuals. The share of unstructured data from social media, videos, and photos is now bigger than the share of structured data.

In addition to the enormous amount of data now generated, stored, and available over networks, the capacity to process the data has increased significantly as well. Complex analyses, involving very large sets of data, are now possible. Data previously seen as unimportant can add critical information when combined with data from other sources. Processing high volumes of a high variety of data at high velocity (real time) is now possible. These three Vs are commonly used to define big data, with value and veracity often added as additional characteristics.

Globally, the big data technology and services market is forecasted to grow about seven times faster than the overall ICT market. According to Gartner, as Figure 1 on the following page indicates, big data is now at the peak of the inflated expectations phase and is forecasted to reach the plateau of productivity in some five to 10 years, a couple of years after cloud computing.

Figure 1: HYPE CYCLE FOR EMERGING TECHNOLOGIES, 2013



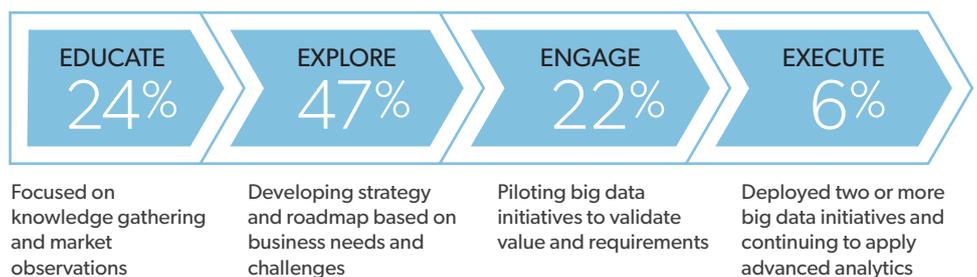
Source: Gartner, July 2013

A number of factors contribute to the current high expectations regarding big data, including:

- The rapid growth of high-speed broadband (both fixed and mobile)
- The increased power of cloud computing
- Easy and cheaper access to big data technologies
- Enhanced connectivity and machine-to-machine communication

Actual usage of big data is nevertheless still relatively low, as evident in a global study showing that only 6 percent of organizations executed big data initiatives in 2012 (see Figure 2).

Figure 2: BIG DATA ADOPTION



Source: Analytics: The Real-World Use of Big Data, a collaborative study by the IBM Institute for Business Value and the Said Business School at the University of Oxford, 2012

## Opportunities Big Data Offers

**“We are big believers in big data.”**

—Participant at the Ministry of  
Information and Communications  
Technology Roundtable

The possibility of making better use of the data already held in an organization can be a major advantage for any institution. The opportunities to create new insights through big data analytics are perhaps most apparent in data-heavy sectors such as healthcare, utilities, transport, education, finance, sciences, and the government. With an enhanced understanding of the people who use their services, these organizations can serve their users better and more efficiently. For example, the Ministry has established a social media research and analysis team. By analyzing large amounts of data from sources such as Twitter, Facebook, blogs, and online forums, the team provides policymakers with timely and comprehensive information on the topics discussed online. Figure 3 below demonstrates the benefits of insights from big data.

**Figure 3: BENEFITS CREATED BY BIG DATA INSIGHTS**



**Big data can be used to identify cost savings and potential efficiency gains through comprehensive analysis of various organizational processes and operations.**

While big data analytics can add significant value to an organization, it is important to identify which processes will provide insights—insights that are adequate, accurate, and actionable—three As to add to the Vs.

Substantial savings in both time and money are expected if organizations and government agencies implement data management practices that consider the specific requirements for big data analysis. Data from different sources inside or outside an organization can bring greater value to multiple functions and for multiple purposes if there is greater transparency. Data collected for a certain purpose might be reused for another purpose, reducing the need for sometimes costly and time-consuming data collection processes.

Big data analysis can also bring a more comprehensive and detailed understanding of different groups' or individuals' characteristics and needs. This is made possible through the inclusion of multiple and differing pieces of data in the analysis. Thus, the services and offerings to these groups or individuals can be more personalized and relevant, and communication and interaction more efficient. As a consequence, the level of satisfaction among customers or residents is likely to increase.

Big data can also make predictive analytics more efficient and accurate. Changes in, for example, preferences and behavior among specific customer groups can be predicted more easily, and organizations can then plan for them more effectively.

Big data can also be used to identify cost savings and potential efficiency gains through comprehensive analysis of various organizational processes and operations. In addition to the productivity gains, the likelihood of identifying opportunities for various innovations can also increase.



## Challenges

---

### Privacy and Trust

**The privacy and confidentiality issues with big data analytics are primary concerns globally as well as in Qatar.**

Access to the cloud, big data technologies, and phenomenal processing power bring new challenges in terms of security as well as privacy. The privacy and confidentiality issues with big data analytics are primary concerns globally as well as in Qatar, as expressed by the participants at the Ministry roundtable. Through analyses of several sources of anonymous data it may, for example, be possible to identify the individuals in the data sets with a high degree of certainty.

Residents', consumers', and organizations' trust that their data will not be misused, accidentally or on purpose, is low, globally as well as in Qatar. Public trust in government agencies and organizations is a prerequisite for Qatar's residents to fully accept that big data analysis can take place without compromising their privacy rights.

Traditionally, enforceable regulations have been at the core of the protection of privacy and integrity rights. With cloud computing and big data, technical approaches to address security and privacy issues are also greatly needed. Such approaches should include secure communication, secure data collection, and secure data storage, as well as granular access control. Approaches that make personal data anonymous can protect the privacy of any particular individual while still providing the needed level of accuracy in the analysis. With high-speed mobile networks and the abundance of smartphones, it may even be possible to offer consumers the option to accept/reject in real time whether their personal data is shared or used. With this opportunity to reject, the benefits to the individual for allowing his/her data to be used will have to be made clear by the organization wanting to use it.

The Ministry is committed to protecting the privacy and integrity of Qatar's public and has drafted the Data and Privacy Protection Law in demonstration of that commitment. Among other Ministry initiatives is the National Information Assurance Framework, which consists of various policy documents and guidelines to address information assurance with respect to existing and emerging technologies, including areas such as governance structure, risk management, security awareness, and incident management.

---

## Access to Data

Big data analysis relies on data that is discoverable, accessible, and usable, and the data must be gathered, stored, utilized, and managed in accordance with all relevant laws and regulations. To reap the full potential of big data, management and government agencies need to focus on making as much data as possible open and available.

The Ministry is publishing an Open Data Policy to facilitate the use of government data within and between government agencies as well as by the public.

Data stored in other countries can be critical for organizations in Qatar as well. At the same time, data stored in Qatar may be of great value to organizations operating outside Qatar. The fact that different countries have different laws and regulations around data storage and how and what data can be accessed provides a significant obstacle for the users of big data.

---

## Data Quality

The data available for big data analysis is by definition abundant and of varying quality. Ministry roundtable participants identified the quality of data collected and shared by using big data as a major area of concern. In particular, the quality of data coming from social networks is seen to be of varying quality.

In the end, improving data quality through the harmonization of APIs (application programming interfaces), data formats, and metadata can further facilitate organizations' and government agencies' use of big data and bring sustainable cost savings and productivity improvements.

---

## Security

Regarding data security, roundtable participants cited the need for a transparent way of classifying data—because it is either critical or sensitive—that must be under absolute control (including physical) by organizations in Qatar. Data owners' consent is needed to move data outside Qatar, but roundtable participants argued that the legal system could have a bigger and clearer role in deciding what data should not leave Qatar.

The Ministry is well aware of the pressing and complex issues around securing big data analysis in a cloud environment and is publishing a cloud security policy as an important component of the National Information Assurance Framework.

---

## Bandwidth

Big data requires reliable and high-speed connections. However, broadband connections with sufficient upload and download speeds were identified by roundtable participants and the Ministry as a major bottleneck for further development of big data in Qatar. Limited bandwidth poses a challenge for accessing and analyzing data inside as well as outside Qatar. One of the major targets of the Qatar National Broadband Plan is “All businesses, schools, hospitals, and government institutions shall have high-quality access to at least 1 Gbps download and upload speeds by 2016.” QNBP also details various policy actions to accelerate and facilitate the rollout of high-speed broadband to meet the target, including acceleration of migration to fiber and improved international bandwidth.

---

## Skills

Big data is a new and complex area, and many organizations need to attract employees with new and diverse skill sets. These include research, statistical, interpretive, and analytical skills as well as ICT skills. In addition, in-depth understanding of the underlying business is required. Developing such comprehensive skills is a challenge to any educational system. Gartner estimates that some 4.4 million big data jobs will be open globally in 2015, and only one-third will be filled.<sup>1</sup>

Roundtable participants agreed that there already is a shortage of big data talent in Qatar and that having skilled staff is more important than having the right software. This is a critical area to address if the promise of big data is to be realized. The need for educational institutions, research bodies, and stakeholders from different sectors to jointly identify requirements for big data experts was underlined. Specific research teams dedicated to big data and the cloud are seen as important as well. Furthermore, as this combination of skills and expertise is hard to find in a single individual, teams of experts will have to work together to get the best value out of big data analytics. For some organizations, this will require both an organizational and a cultural shift.

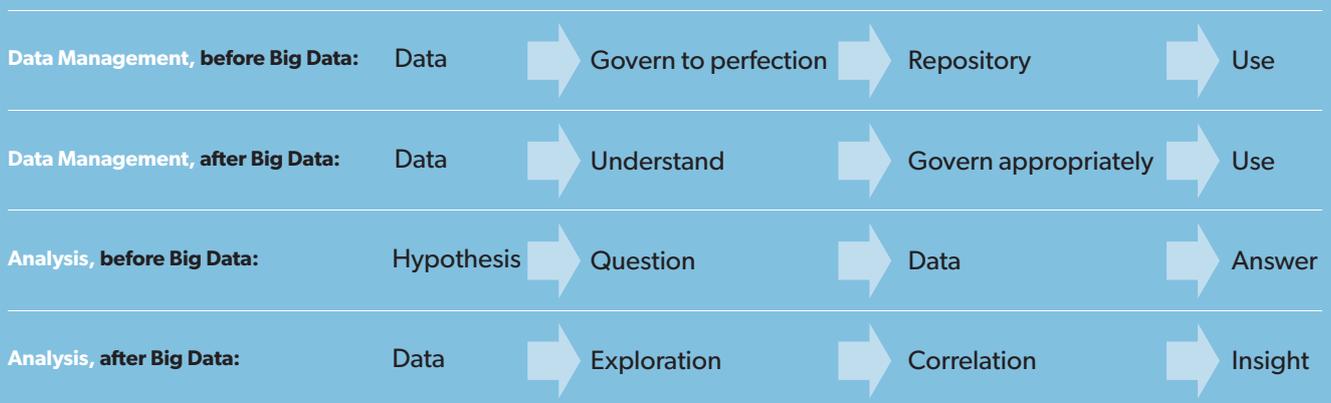
<sup>1</sup> Gartner, “Gartner Reveals Top Predictions for IT Organizations and Users for 2013 and Beyond,” press release, October 24, 2012.

## Future

The advantages from being able to scan through large amounts of data and combine data from several sources to provide a more comprehensive analysis are clear. The rollout of ubiquitous and reliable high-speed connections will enable widespread adoption of big data in Qatar. Decision makers and policymakers in all sectors of Qatar are likely to be relying more and more on insights generated from big data analytics.

With maturing big data and cloud computing technologies, the focus will shift from how to store and process large amounts of data to how to make informed decisions in timely, accurate, and cost-efficient ways. This may require a paradigm shift in how data is managed and analyzed, as seen in figure below.

**Figure 4: DATA MANAGEMENT AND ANALYSIS, BEFORE AND AFTER BIG DATA**





## Conclusion

As this paper indicates, big data holds great promise for all sectors in Qatar. However, public trust in government agencies and organizations is required before the people of Qatar will fully accept and adopt big data analysis. Organizations and government agencies need to demonstrate that they are not compromising the privacy and integrity of the population in Qatar when implementing big data. The Data and Privacy Protection Law needs to be implemented and acted upon to demonstrate that the government is properly addressing privacy issues. In addition, technical solutions should be used where possible.

The implementation of the National Information Assurance Framework policies and guidelines is key to the adoption of big data in Qatar. Legislation is needed, for example, to clarify which data can go to a cloud environment or leave Qatar, and which data cannot. Clearer rules for accessing data stored in Qatar from abroad are needed as well. An analysis of the various types of data generated and available in Qatar must be executed before such a classification of data is done.

To stimulate big data-enabled innovation, government agencies can lead by example and release public sector data as per the Ministry's Open Data Policy.

The global and local shortage of big data experts and skilled users is a serious challenge to big data adoption in Qatar. Educational institutions and industry need to jointly define the required skills for big data and offer specific big data education. Collaboration between universities and companies, with students being involved in the corporations' big data projects and activities, is recommended. On-the-job training and other forms of continued education are needed as well.



**We would like to thank the organizations below for their active participation and contributions during the Ministry of Information and Communications Technology Big Data Roundtable.**

- |                              |   |   |
|------------------------------|---|---|
| Cisco Systems                | Ministry of Interior                                | Qatar National Bank (QNB)                 |
| Gulf Business Machines (GBM) | National Institute of Standards and Technology, USA | Qatar Petroleum                           |
| IBM                          | Ooredoo   | Qatar University                          |
| Q-CERT                       | Qatar Airways                                       | Supreme Council of Health (written input) |
| Kahramaa                     | Qatar Computing Research Institute                  | Vodafone                                  |
| MEEZA                        |   |   |

The information contained in this paper is for general information purposes only. The information is provided by the Ministry of Information and Communications Technology, based on desk research and roundtable discussions. The Ministry endeavors to provide up-to-date and correct information, but makes no representations or warranties of any kind, expressed or implied, about the completeness, accuracy, reliability, or suitability of the information in the paper, for any purpose. Any reliance placed on such information is strictly at the reader's own risk. In no event will the Ministry be liable for any loss or damage including, without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from loss of data or profits arising out of, or in connection with, the use of this paper.



## References

*National Broadband Plan for the State of Qatar*, 2013; <http://www.ictqatar.qa/en/documents/document/qatar%E2%80%99s-national-broadband-plan>.

Gartner, “*Gartner Reveals Top Predictions for IT Organizations and Users for 2013 and Beyond*,” press release, October 24, 2012; <http://www.gartner.com/it/page.jsp?id=2211115>.

IBM Institute for Business Value and the Said Business School at the University of Oxford, *Analytics: The Real-World Use of Big Data*, 2012 (updated July 29, 2013); [http://www.ibm.com/smarterplanet/global/files/se\\_\\_sv\\_se\\_\\_intelligence\\_\\_Analytics\\_-\\_The\\_real-world\\_use\\_of\\_big\\_data.pdf](http://www.ibm.com/smarterplanet/global/files/se__sv_se__intelligence__Analytics_-_The_real-world_use_of_big_data.pdf).

Ericsson White Paper, *Big Data Analytics*, August 2013; <http://www.ericsson.com/res/docs/whitepapers/wp-big-data.pdf>.