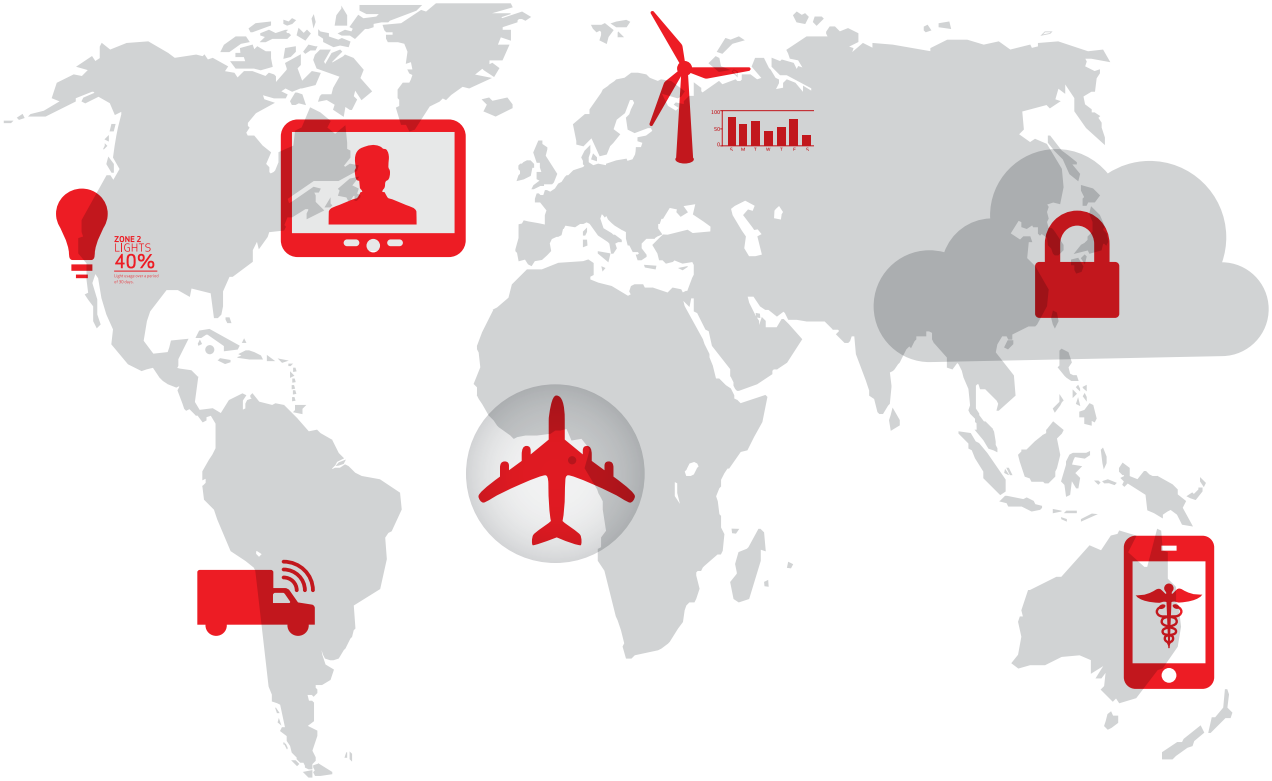
How to thrive on the FRONTIERS OF DATA

What big data means for you, and how cloud can help you become a data-driven business



How to thrive on
THE FRONTIERS OF DATA

Finding effective strategies for leveraging big data—ways to gather the right data, process it and extract valuable insight—is fast moving up the list of corporate priorities.

Contents

BIG DATA: HYPE OR REALITY?

According to a study by the Economist Intelligence Unit:[1]

- Nine out of ten business leaders believe data is now the fourth factor of production, as fundamental to business as land, labor, and capital.
- The use of big data has improved business performance by 26% and its impact will grow to 41% over the next three years.
- 58% of companies say they plan to increase investment in big data over the next three years.

According to Gartner:

- 42% of IT leaders stated they had invested in big-data technology, or were planning to do so within a year.[2]
- Organizations that integrate high-value, diverse information types and sources into a coherent information management infrastructure will outperform their industry peers by more than 20%.[3]

Companies that harness the intelligence inherent in their data, secure it, and act on it accordingly, can gain a significant competitive edge.

But Gartner also notes that big data is "falling into the trough of disillusionment"[4]—the beginning of a difficult process of maturing toward becoming business as usual. Enterprises are starting to ask tough questions about what big data can actually deliver for them.

And that's what this paper is about. We look at what big data can offer, and how you can get started with your own big-data initiatives, using technologies such as cloud to help you make the step-change to become a data-driven business.

Defining big data

THE FOUR Vs

Most definitions of big data agree that it involves the "three Vs": volume, velocity, and variety. We add a fourth: value. This is in many ways the most important.

THE ACCEPTED THREE VS

When defining big data it's common to discuss three Vs: volume, velocity and variety.

Volume: Although there's no fixed number marking the beginning of "big", we're talking much bigger than conventional tools like spreadsheets and relational databases can handle easily. Many case studies of big data involve datasets of many petabytes—or even exabytes—made possible only by using high-performance cloud-based computing.

Velocity: Many big-data applications, such as cancer research, use historical data, but much attention is being paid to how to leverage real-time data—not just collected in real time, but processed and accessed in real time too. In many scenarios, users must be able to ask questions iteratively and get answers in minutes, not days.

Variety: Big data covers not just "structured" data neatly normalized into a fixed schema and exported from ERP or CRM systems. It also includes semi-structured data, (which, although it has no fixed configuration, is categorized using tags or other metadata) and unstructured data, such as email messages and videos.

Not all these three Vs must apply for data to be called big data. For example, data might need to be analyzed in real time, but be relatively small in volume and reside largely in structured formats.

THE FOURTH V: VALUE

What makes big data exciting is not the scale and diversity of the data, but the scale and diversity of the things you can do with it. That's why we add a fourth V.

Value: Big data can generate business value by supporting specific tactical decisions across the business, generating more accurate forecasts and identifying trends, or even automating processes such as pricing or stock ordering.

> "Big data is about looking ahead, beyond what everybody else sees."[6]
>
> Peter Sondergaard, senior vice president and global head of research at Gartner

Descriptive analytics is the type of business intelligence that we're all familiar with. Analysts run queries against the data warehouse to answer questions about what's already happened—how many widgets did we sell last month versus a year ago.

The main focus of most big-data projects is predictive analytics. We don't necessarily need to know how to ask a question or which data items we need to query; instead, we rely on algorithms that find answers in very large data stores. We may want to find out:

- What characteristics should we be looking for in new customers?
- What patterns are there in our sales volumes that would help us model production and set stock levels?

And with the increasing maturity of open-source big-data tools driving down the cost of processing data, the potential of big data is now accessible to nearly any enterprise.

70%

A SURVEY FOUND THAT 70% OF RESPONDENTS CAN ENVISION A **"KILLER APPLICATION"** *FOR BIG DATA THAT WOULD BE* **"VERY USEFUL"** *OR* **"SPECTACULAR"** *FOR THEIR BUSINESS.*[5]

Definition of Big Data:

Big data is data that—given its volume, velocity, variety, and complexity—requires new techniques and technologies to make extracting its value at an extreme scale economical.

Six trends that are
DRIVING BIG DATA

The demand for better data to base decisions on has always been there, and business intelligence has been improving for some time. But in the last few years six trends have driven a step change in what's possible.

80%

EXPERTS ESTIMATE THAT AROUND 80% OF ENTERPRISE OR TRANSACTIONAL DATA HAS SOME LOCATION-RELATED COMPONENT.[7]

EMERGENCE OF POWERFUL COMPUTING 'IN THE CLOUD'

The availability of rapidly scalable storage and computing resources, without massive upfront investment in in-house infrastructure or staffing, has transformed the economic and technological possibilities for processing vast amounts of data.

Advances in technologies, specifically scalable distributed computing and storage frameworks, have enabled data scientists to build models with massive data sets—utilizing entire populations instead of just sampling—to achieve new levels of accuracy.

INCREASING PERVASIVENESS OF CONNECTIVITY

The near ubiquity of networks closes the loop, giving individual users—whether they're field engineers, shop-floor workers, or the CEO—instant access to information literally in the palm of their hands practically wherever they are.

IT-SAVVY EMPLOYEES WORKING OUTSIDE IT

IT skills used to be concentrated in the IT department, but not anymore. The children of Gen Y are now entering their thirties, and make up a sizable proportion of management roles in all lines of business. They've grown up with IT and demand more from it. And if the IT department can't help, they'll look elsewhere or do it themselves.

GROWING VOLUME OF THIRD-PARTY DATA

Both public and enterprise social networks are producing a "fire hose" of real-time textual, image, video, and file data, complete with metadata about relationships, context, and connections. Many governments and supranational organizations (like the EU) are making data—including mapping, weather and economic data—freely available in the cloud.

AVAILABILITY OF SMALL, LOW-COST SENSORS

Billions of tiny machine to machine (M2M) sensors are now sending a constant stream of data—for example, vehicles transmitting their location; smart meters sending details of energy use; or industrial equipment reporting its operational status. All this data is not only of direct use to operational processes (like billing customers), but also builds a wider picture that's relevant to many other activities, from forecasting demand to designing new products.

PROLIFERATION OF GEOLOCATION DATA

Many experts estimate that around 80% of enterprise or transactional data has some kind of location-related component,[7] such as an address or coordinates, which provides valuable context. Building this location data into big-data analysis can reveal correlations between different data sets, and produce unique insights.

The big data LIFECYCLE

Big data isn't just more of the same. It's a fundamental shift in how enterprises see and use information. It requires a step change at each stage of how data is generated, gathered, processed, and acted upon.

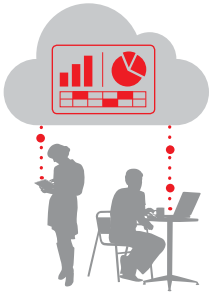
THE BIG-DATA LIFECYCLE COMPRISES FOUR STAGES:

Acquisition
The collection of data from a wide range of sources

Marshaling
The organization and storage of data

Analysis
Finding insights, modeling, making predictions

Action
Changing business outcomes



- **Collection/Acquisition:** Data from internal and external sources, owned or bought, in a variety of formats, to meet identified business needs. Data can exist in many disparate stores and across several domains, and may need special processing during collection.
- **Aggregation/Marshaling/Orchestration:** Organizing and storing data ready to be analyzed according to governance standards. Big data provides a complete centralized view of disparate data.
- **Processing/Analysis:** Including predictive behavioral modeling, pattern recognition and summary, producing insights to change business outcomes. Big-data analytics is often performed in cloud-based environments as they allow easy scaling of storage, give access to massive on-demand computing power and are isolated from "business as usual" workloads.
- **Sharing, Decision, Access, and Action:** Getting the data to the right place so it can be used for a business advantage, through dashboards, data services, and reports to PCs and mobile devices. Cloud computing and web technologies have made it much easier to deliver information and tools to users wherever they are and whatever device they are using. And application programming interfaces (APIs) and services allow outputs from big-data analyses to be integrated into existing systems.

IN ORDER TO REALIZE THE POTENTIAL OF BIG DATA, YOU NEED TO ENSURE THAT THE RIGHT INSIGHT IS AVAILABLE TO THE RIGHT PEOPLE WHEREVER THEY ARE.

Opportunities across industries

BIG DATA IN ACTION

These examples illustrate the huge variety of applications for big data. Some produce a real competitive edge; others incremental improvements in operational efficiency. The possibilities are almost endless.

30%

INTERCONTINENTAL HOTELS USED BIG DATA TO HELP **INCREASE CONVERSION BY OVER 30%.**

RETAIL AND HOSPITALITY: INCREASING ENGAGEMENT

The better you understand your customers and their needs, the more you'll be able to develop the right products, set the right stock levels, tailor your marketing, improve your customer service, identify trends and forecast demand, and ultimately outflank your competitors.

Big data can help by crunching through the fuzzy content of social networks—particularly brand sentiment—as well as data from sales and stock systems, CRM, and loyalty programs. Finding correlations in buyer behavior and demographics can help build personalized profiles and dramatically improve customer engagement.

Want to improve customer engagement? Big data can help build much more detailed buyer profiles and increase conversion.

"InterContinental Hotels Group consolidated information about its 71 million Priority Club customers and social media data into a single data warehouse. Using the system, it launched a new marketing campaign. Where previous campaigns might have seven to 15 customized marketing messages, the new campaign has 1,552, each defined by 4,000 attributes. The campaign has increased customer conversions by 31%."[8]

GOVERNMENT: REACHING THE INDIVIDUAL

The 2012 U.S. presidential election has been called the first "big data" campaign. Like a retailer using data to profile customers, the Democratic campaign micro-targeted voters based on their individual preferences. Using data from external lists, historical voting behavior and surveys, the campaign adjusted its approach to individuals in real time.

And big data continues to add value after election. Government and law enforcement can monitor social media to help identify security threats, detect fraud, track the popularity of policies, and even monitor the flow of people within and between countries.

FINANCIAL SERVICES: CALCULATING RISK

Insurers of all kinds are using customer data to detect fraud and to better profile and model risk, to set more accurate premiums. For example, car insurers, in partnership with Verizon Telematics, can now use M2M data gathered from vehicle sensors—such as speed, miles traveled, and even things like harshness of braking—to score individual driving behavior and set personalized premiums.

Financial services firms can also much more quickly assess and forecast risk, whether that's for an individual's credit score or for an overall investment portfolio.

R&D: ACCELERATING INNOVATION

Organizations focused on research, already heavy users of high-performance computing and large data sets, are seeing some of the biggest payoffs from big data. For instance, in genomics research, advances in sequencing technology have produced vast amounts of data by making the research faster and far less expensive. Consumer-oriented business models in genomics can also help speed medical research. Take, for example, MIT and Harvard's Human Genome Project at the Broad Institute. This substantial research "yielded [a] permanent foundation for biological research, and launched a new era in biomedicine." [9]

Organizations focused on research are seeing some of the biggest payoffs from the use of big data.

"NASA is involved in analyzing data collected from planes to study safety implications, which in turn will help with commercial airlines' maintenance procedure improvements and potentially prevent equipment failures. Using advanced algorithms, the agency helped tease out relevant information from a mountain of unstructured data to help predict and prevent safety problems. Using the open-source Multiple Kernel Anomaly Detection (MKAD) algorithm, the agency determined how two continuous data streams or networks were similar, and then analyzed them using a single framework to detect patterns to automatically discover precursors related to adverse events while an airplane is in flight."[10]

HEALTHCARE: HELPING IMPROVE OUTCOMES

Healthcare and life sciences organizations are some of the leading users of big data. It's been used to predict the progress of flu epidemics by mining data from social networks; to forecast health risks when combining different environments, lifestyles, and genetic environments; to analyze clinical trial results; and to help improve individual patient care by bringing together insights drawn from electronic medical records and data gathered by connected patient monitors. Shared cloud platforms can help organizations around the world collaborate and pool knowledge for better outcomes.

Verizon is working on a project with NantWorks to help develop the Cancer Knowledge Action Network (CKAN), a cloud-based database system that will help doctors provide efficient and effective cancer care through better decision-making and improved productivity.

NantWorks and Verizon are working on translating big data into information accessible for cancer treatment.

ENERGY AND UTILITIES: PLANNING FOR THE FUTURE

Long aware of the value of data, energy firms are already using proprietary data to model where to drill for oil or gas, and combining internal data with data about stock prices, shipping costs, and customer demand to build better forecasts.

Verizon is working with one utility company to collect data from thousands of smart meters and electric-vehicle charging stations. Data about energy usage is captured every 15 minutes, and shared via wireless links to the cloud where it's processed and used to help manage load control.

Five steps for
BIG DATA SUCCESS

So you're excited by the results that other organizations are seeing from big data. How do you get in on the action? There are several considerations, covering people, process and technology.

Gartner predicts that only one-third of the 4.4 million IT jobs to be created by big data by 2015 will be filled, noting that: ***"Data experts will be a scarce, valuable commodity."*** [11]

1. TAKE ADVANTAGE OF CLOUD

Big-data workloads can place a significant burden on your infrastructure. You'll need processing power to crunch the data sets quickly, storage to hold data at rest, networks to move it around, and analytics applications to manage it and give users access to results.

Many businesses find that expanding their own data center facilities is too big a drain on budgets and too long a commitment, considering that you'll still be at an experimental stage with big data. What's more, it's hard to predict what infrastructure you'll need.

That's why Infrastructure as a Service (IaaS) offers an ideal platform for big data. It gives quick access to sizeable resources, and the ability to scale capacity as your needs grow, without upfront capital investment in infrastructure, facilities or IT administrative personnel.

Infrastructure as a Service is the ideal platform for big-data initiatives.

You'll probably also need specialists to model the data, work the analytics tools, and build the algorithms. This team should work closely with internal subject matter experts and draw on their deep knowledge. Because big-data initiatives have become so popular so quickly, some of the necessary skills are in short supply and the best people can command a premium.

Given all this change, you're likely to consider bringing in a third-party. If so, it's critically important to evaluate the security of their operations, and quiz them about data portability and openness. The last thing you want to find is that you're locked into a provider and can't easily move your data.

2. BUILD A COMPELLING BUSINESS CASE

We recommend that you pay close attention to your business objectives and build a business case. Big data is not an IT-driven solution — just investing in more infrastructure to ingest and process data won't improve business results. The question to ask is: ***"What could we do with big data that we can't do without it?"***

Begin by identifying the problem you want to solve within the context of a particular process. Are you trying to prove a hypothesis or explore a field? Perhaps it's understanding your customers better to feed into product development, or identifying waste in your supply-chain processes. That's not to say you won't find some unexpected ways to apply the data you already have, but like all initiatives, your big-data efforts should begin with some defined use cases. Practically speaking, that means big data becomes a joint effort between the lines of business and IT.

Engage stakeholders from across the business when building the business case.

Using big data represents a different way of thinking, one that embraces change and uncertainty instead of approaching problems with assumptions and preconceived ideas. Some managers will resist, seeing big data as a threat. You'll need to tackle these fears head-on, and the best way to do so is to involve the wider business as much as possible when building your business case.

3. TACKLE YOUR DATA ENVIRONMENT

What data do you need to meet your business goals? Taking an inventory of your data sources is a critical step. Many organizations find that they have silos of data scattered around the business—you may even have the same data stored in multiple locations for multiple purposes, owned by different groups, such as financial auditing and sales forecasting.

As Gartner notes, you'll need to ***"shine a light on 'dark data'"***[11]—the valuable data you're collecting, which may be going unused. The advent of big-data technologies can expose insights that traditional business intelligence tools cannot.

Most enterprises aren't leveraging external data to any real extent, so there's a genuine opportunity to get ahead.

You can also start using big-data technologies on newly gathered data, and source or buy new information externally. This is where imagination comes in to play. Look beyond your ERP and CRM systems; data may flow from M2M sensors, social media feeds, documents on your network, transactions, or operational systems, as well as images, audio recordings, videos, location data, third-party feeds about everything from the weather to stock market performance, and more.

4. ADAPT YOUR GOVERNANCE AND SECURITY MODELS

Big data can put a massive strain on your information governance processes. Every bit of data you gather must be assessed for quality, stored, organized, tracked and audited, described with metadata, de-duplicated, secured, backed up, analyzed, and contextualized. Some of the data will be commercially sensitive and some personal—like medical records. And data may be spread much more widely, not constrained within the familiar silos of functional systems.

As a result, you must plan for both security and compliance very carefully. Not only should you consider how to govern access to and exposure of data, and how to protect data at rest and in transit, you'll need to consider how you can audit your access controls for information security compliance, and manage issues such as data retention and residency—you may be required to store some data within national boundaries, or only retain certain data for a limited time.

The goal is to achieve an appropriate balance of availability and security.

Having the tools to interpret, secure, and manage data based on its value, and achieving an appropriate balance of availability and security, has become a business-critical corporate goal.

5. CHANGE RIGHT ACROSS THE BUSINESS

While the governance of data assets and supporting infrastructure is a major consideration, the most important impact of big data should be on your business processes. Everyone in the business—from marketers planning the next quarter's campaign to call-center managers deciding on the next week's staffing levels—must be cognizant of the potential that data-driven insight can offer, anad not just be carrying on with business as usual.

To capture data, you may have to change front-line processes and how they feed into central decision making. To get actionable results out of the insights your analytics systems produce, you need to distribute findings quickly to managers and staff across the business, particularly where decisions get made: in stores, field marketing departments, and warehouses.

Insight is worthless if the business can't make timely use of it.

Data is worthless if the business can't make use of its insights in time. The goal is to create an empowered, real-time business where employees can connect to cloud-based dashboards or analytics tools—even when mobile—to get access to the insight or alerts they need.

*To capture data, you **may have to change front-line processes** and how they feed into central decision-making.*

BECOMING DATA-DRIVEN

Big data is an initiative that involves the entire business. But the IT function has a critical role to play. Keeping cost and risk under control is vital.

*Big data can take corporations out of their comfort zone, revealing new insights that **challenge the status quo.***

FOCUS ON THE DATA, NOT THE INFRASTRUCTURE

Big data doesn't just require more processing power, it also places a burden on other parts of your infrastructure—gathering and transmitting data requires bandwidth and the data has to be stored and backed up.

When you're focusing on challenging new areas like analytics, you don't want to be spending your time worrying about infrastructure. That's why Infrastructure as a Service (IaaS) is ideal for big data. It gives access to a vast amount of processing power without upfront investment.

Infrastructure as a Service provides the ideal platform for big data, giving you cost-effective access to a vast amount of storage and compute resources.

Storage as a Service (STaaS) is also a great fit for big data. Not only is it cost-effective, it also enables you to limit the burden on existing production systems—you won't want your forays into big data to affect order processing, for example.

TEST THE WATER WITH A PILOT PROJECT

It can be hard to understand the potential of big data, or what it takes to make an initiative happen, without actually trying it out for yourself. But you can't dive in at the deep end. Pick a pilot project. You might start, for instance, applying big-data technologies and platforms to get new insights from your existing 'dark data', experimenting along the way to gain experience and knowledge. Or you might apply big-data technologies to new streams of gathered data, to assess their value and the questions they might be able to answer. Cloud can help you here, too, giving you a test environment separate from your production systems.

TAKE THE LEAD ON GOVERNANCE

For IT, a major focus area has to be how you govern and secure information assets. And this will include pinning down what the business plans to do with data, why, and what the limitations should be. Some big-data initiatives have backfired, for example, because consumers have perceived large-scale data gathering as an invasion of privacy. These risks need close consideration, and communication is vital.

CHOOSE PARTNERS CAREFULLY

Big data is complex, new and still evolving quickly, so choosing an experienced partner is an important step in your journey. Above all, you need a partner that knows the value of data firsthand in all its forms. Beyond that, look for one that appreciates the unique needs of your business, and understands how big data fits into the other strategic initiatives you're working on today—such as cloud, mobility and M2M.

ABOUT VERIZON

As well as driving development of big-data solutions and the infrastructure that powers them, we have been actively using big data to improve our own business for many years.

HOW WE USE BIG DATA

Managing nearly 5,000 networks and more than 340,000 security, network and hosting devices in more than 150 countries generates a lot of data. We use that data to improve the performance, reliability and security of our global infrastructure and to shape our investment plans so that we have the capacity to meet our customers' needs.

Examples include:

- We use the vast quantities of data collected from our hundreds of thousands of network routers, switches, firewalls, and data centers to plan future capacity needs and provide early warning of potential problems so we can take preventative action.
- We leverage mobile device location information to keep our field workers safe and optimize their routes and schedules to provide better service levels for our customers.
- We analyze data on network security attacks to build sophisticated profiles of attack signatures—we then provide this as a feed to our customers, helping them to improve their defenses.

HOW WE CAN HELP YOU

Cloud: Converting big data into actionable intelligence requires massive computing power, but often only for short periods. Our cloud-based solutions give you access to vast storage and compute resources without huge capital investment. And you only pay for what you need.

Networking: Our pervasive networks offer cost-effective, reliable and secure ways to gather data and disseminate insight to those that need it. From fixed line to powerful 4G LTE wireless and satellite, we can help you connect sensors and workers.

M2M: One of the biggest areas of growth in big data is embedded sensors. Cisco predicts that by 2020, the number of things connected to the internet will exceed 50 billion—from cars to pacemakers. Verizon's M2M platform provides a secure way to connect and manage the data from all your assets, enabling you to focus on uncovering insight and providing tools for your employees.

Getting started: Embracing big data may seem like a daunting prospect, but it needn't be. Our Big Data Maturity Assessment provides a baseline of where you stand and how you can grow, addressing all the technical and organizational challenges we've discussed. It puts those challenges into perspective and helps you build an action plan.

We're an active participant in the standards bodies that are working to make big data more robust, more secure and ***more powerful.***

To find out more about our solutions for big data, and to talk to us about your needs, contact your account manager or visit verizonenterprise.com

References

1 uk.capgemini.com/resources/the-deciding-factor-big-data-decision-making
2 gartner.com/newsroom/id/2366515
3 sas.com/offices/NA/canada/lp/Big-Data/Extreme-Information-Management.pdf
4 blogs.gartner.com/svetlana-sicular/guest-post-big-data-is-falling-into-the-trough-of-disillusionment/
5 aiim.org/pdfdocuments/iw_big-data_2012.pdf
6 gartner.com/newsroom/id/2207915
7 critigen.com/blog/what-is-location-analytics and gislounge.com/80-percent-data-is-geographic/
8 online.wsj.com/article/SB10001424127887324178904578340071261396666.html
9 broadinstitute.org/scientific-community/science/projects/mammals-models/human/human-genome-project
10 open.nasa.gov/blog/2012/10/04/what-is-nasa-doing-with-big-data-today/
11 gartner.com/newsroom/id/2207915
12 share.cisco.com/internet-of-things.html

This report would not have been possible without the contributions of Mike Palmer, VP, Verizon Product Strategy; Tonnetta Oubari, Executive Consultant, Verizon Global Professional Services; Wilco Van Ginkel, Verizon Enterprise Strategy and Co-chair of the CSA Big Data Working Group; and Ashok N. Srivastava, PhD, Verizon Chief Data Scientist.

verizonenterprise.com