Insights on governance, risk and compliance

April 2014

Big data

Changing the way businesses compete and operate

EY

Building a better working world

Contents

Big opportunities, big challenges

The idea of data creating business value is not new, however, the effective use of data is becoming the basis of competition.

Business has always wanted to derive insights from information in order to make better, smarter, real time, fact-based decisions: it is this demand for depth of knowledge that has fueled the growth of big data tools and platforms.

Those leading the change are now including big data from both within and outside the enterprise, including structured and unstructured data, machine data, and online and mobile data to supplement their organizational data and provide the basis for historical and forward-looking (statistical and predictive) views.

Big data will fundamentally change the way businesses compete and operate. Companies that invest in and successfully derive value from their data will have a distinct advantage over their competitors — a performance gap that will continue to grow as more relevant data is generated, emerging technologies and digital channels offer better acquisition and delivery mechanisms, and the technologies that enable faster, easier data analysis continue to develop.

While the ability to capture and store vast amounts of data has grown at an unprecedented rate, the technical capacity to aggregate and analyze these disparate volumes of information is only just now catching up.

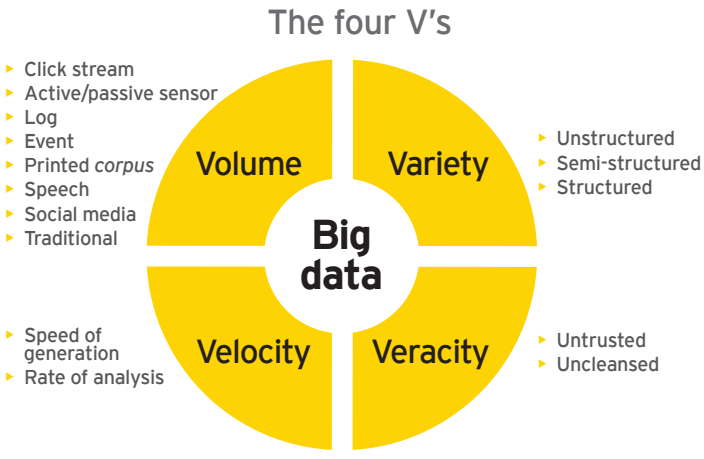
This report explains what big data is and how you can benefit from using it in your business operations, especially in the field of analytics. Using our in-depth client knowledge, we outline some of the opportunities already grasped by global organizations. In addition, we also highlight some of the key risks — some you may already be aware of, but others that may be new to you when you join the big data revolution.

What is big data?

Big data refers to the dynamic, large and disparate volumes of data being created by people, tools and machines; it requires new, innovative and scalable technology to collect, host and analytically process the vast amount of data gathered in order to derive real-time business insights that relate to consumers, risk, profit, performance, productivity management and enhanced shareholder value.

Big data includes information garnered from social media, data from internet-enabled devices (including smartphones and tablets), machine data, video and voice recordings, and the continued preservation and logging of structured and unstructured data. It is typically characterized by the four "V's":

- **Volume:** the amount of data being created is vast compared to traditional data sources
- **Variety:** data comes from different sources and is being created by machines as well as people
- **Velocity:** data is being generated extremely fast – a process that never stops, even while we sleep
- **Veracity:** big data is sourced from many different places, as a result you need to test the veracity/quality of the data



Evolving technology has brought data analysis out of IT backrooms, and extended the potential of using data-driven results into every facet of an organization. However, while advances in software and hardware have enabled the age of big data, technology is not the only consideration. Companies need to take a holistic view that recognizes that success is built upon the integration of people, process, technology and data; this means being able to incorporate data into their business routines, their strategy and their daily operations.

Organizations must understand what insights they need in order to make good strategic and operational decisions. The first part of the challenge is sorting through all of the available data to identify trends and correlations that will drive beneficial changes in business behavior. The next step is enriching this organizational information with that from sources outside the enterprise; this will include familiar big data sources, such as those created and stored online.

In a business environment that constantly and rapidly changes, future prediction becomes more important than the simple visualization of historical or current perspectives. For effective future prediction, data analysis using statistical and predictive modeling techniques may be applied to enhance and support the organization's business strategy. The collection and aggregation of big data, and other information from outside the enterprise, enables the business to develop their own analytic capacity and capability, which for many years has only been available to a few larger organizations.

The impact of big data

To understand the impact of how data has transformed our daily lives, look no further than how the movie rental experience has changed. When movies were rented from independent neighborhood stores, the rental agent would base their recommendations on which movies the customer said they liked and a large amount of their own opinion.

Today, movie rental companies and content delivery services can utilize a vast array of data points to generate recommendations. By analyzing what was viewed, when, on what device (and even whether the content was fast forwarded, rewound or paused), as well as user activities such as internet searches, and browsing and scrolling within a webpage, recommendations can be tailored for millions of customers in real time andapproximately 75% of views at a leading provider are now driven by these recommendations.

Technology megatrends

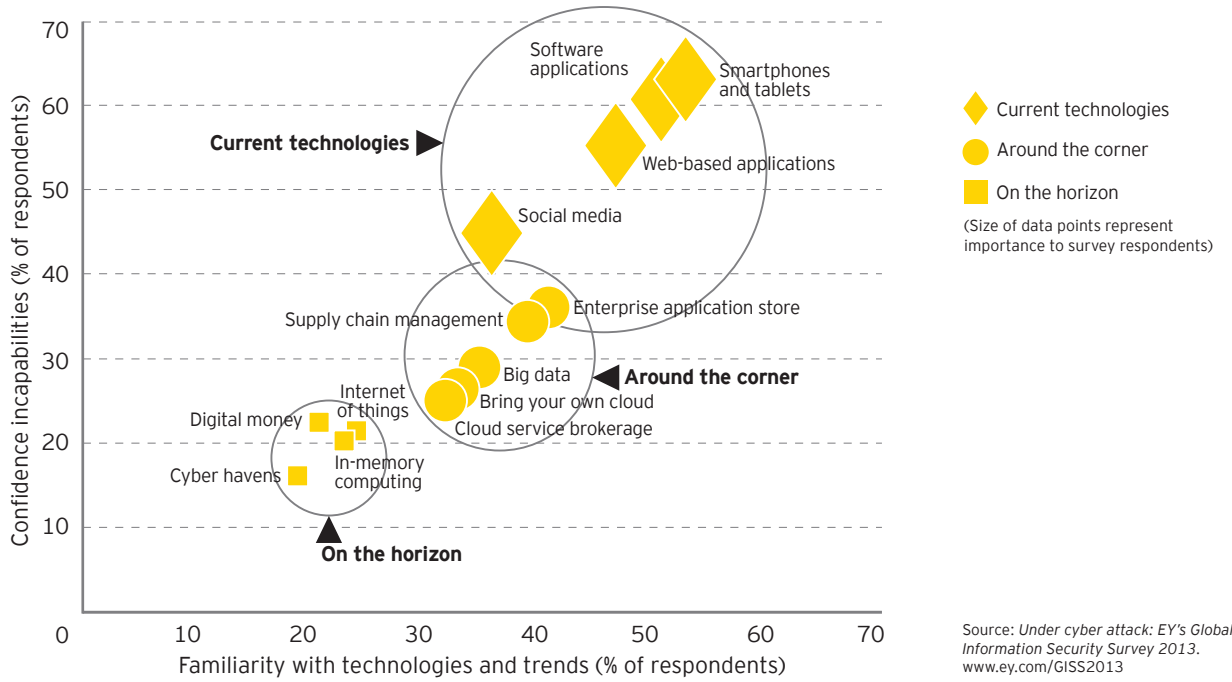
Big data is generating an intense amount of attention among businesses, media and even consumers, along with analytics, cloud-based technologies, digital channels and data visualization. These are all part of the current diverse ecosystem created by the technology megatrends. Some even herald the potential transformative power of the current trends as rivaling that of the internet. Yet, as in the early days of the internet, there is uncertainty about just what big data is, its potential benefits and the associated risks.

EY's 2013 Global Information Security Survey results indicate that while adoption and use of big data is not yet widespread, there is growing confidence and familiarity with the technology. Respondents ranked big data technologies as being "around the corner" (i.e., those that have been on organizations' radar for a period of time but may not yet be implemented or widely adopted) as average in terms of level of importance, familiarity and confidence in their capabilities to address related cyber risks. Organizations typically view these technologies as offering opportunities to improve their performance and create competitive advantage. This is where familiarity and confidence in capabilities needs to increase today, as the importance of these technologies is likely to grow significantly in the near future.

The term Big Data has become a major theme of the technology media, but it has also increasingly made its way into many compliance, internal audit and fraud risk management-related discussions. In EY's *Global Forensic Data Analytics Survey 2014*, 72% of respondents believe that emerging big data technologies can play a key role in fraud prevention and detection. Yet only 7% of respondents were aware of any specific big data technologies, and only 2% were actually using them.

Forensic data analytics (FDA) technologies are available to help companies keep pace with increasing data volumes, as well as business and regulatory complexities; examples can include real-time analytical processing engines that make rapid business decisions, such as stopping a potentially improper payment or business transaction, or leveraging anti-fraud/anti-corruption monitoring controls that integrate data visualization, statistical analysis and text mining. Yet despite their availability, many companies have not scaled up their data usage to take advantage of these effective tools, and may be missing important fraud prevention and detection opportunities by not mining larger data sets to more robustly monitor business activities.

Emerging technologies and trends



Source: *Under cyber attack: EY's Global Information Security Survey 2013*. www.ey.com/GISS2013

Big data eliminates intuition

Decisions can be made with a structured approach through data-driven insight, including:

- Customer and product profitability
- Customer acquisition and retention strategies
- Customer satisfaction strategies
- Marketing segmentation
- Operations and performance management
- Supply chain and delivery channel strategy

Big data life cycle

Creation

Certain types of data have long been able to be captured, but this data has rarely been used effectively until now (e.g., the location of a person at any point in time, the number of steps a person takes every day, a real-time history of credit card purchases). New technology such as advanced sensors and customized software can now record this information for analysis.

Changes in the way we communicate (e.g., social media vs. telephone vs. text/SMS vs. email vs. letter) have also increased our ability to investigate areas such as consumer sentiment. Social media increases the speed at which data is generated; for example, a product launch that is discussed live on a popular social networking site can generate a buzz in real-time and allow companies to gauge public reaction even before the launch event is over.

Processing

Extremely large volumes of data have traditionally not been captured and processed for various reasons, most notably because the cost to do so was far greater than the value of insights companies could derive from its analysis. However, multiple factors and new technologies have lowered the cost and technology barrier for effective data processing, allowing companies of all sizes, to be able to unlock the value contained in different data sources. For instance, it is difficult for conventional relational databases to handle unstructured data, so software frameworks like Hadoop(R), for distributed storage and parallel processing of large datasets have been introduced to process non-structured data at high speed; making it easier to perform a more comprehensive analysis of big data.

Many organizations are looking to the cloud to provide a storage solution that is agile and enables unparalleled scalability; however, these organizations need to ensure the governance and risk management practices on their cloud are appropriate for the type of information being collected. Cloud computing enables companies to use prebuilt big data solutions, or quickly build and deploy a powerful array of servers, without the substantial costs involved in owning physical hardware.

Output

Although it is now easier and cheaper to capture, store and process data, it is not useful unless the information is relevant; it must also be readily available to the right people who need the appropriate input in order to make insightful decisions leading to successful outcomes.

There are three key enablers:

- **Mobile** – established mobile networks have allowed for easier distribution of information in real-time
- **Visual/interactive** – technologies have brought the ability to review large and complex data sets into the realm of the average business user
- **Human resource** – there is a new breed of employees with the knowledge to handle the complexities of big data and with the ability to simplify the output for daily use

Resources and processes

An important factor in being able to achieve big data success is having knowledgeable and competent resources. This extends beyond the so-called data scientists who have deep knowledge and experience in handling, analyzing and reporting on big data sets. While these skillsets are indeed in high demand, success requires more than having a handful of specialists on the workforce.

While governments continue to call for the training of "data scientists," companies are taking advantage of fundamental skillsets that already exist within their organization. Employees with the curiosity to ask the right questions and the ability to synthesize and leverage new data points quickly are well suited to lead the big data revolution. In reality, *they are the revolution,* but they must be supported with business processes that place value on gathering and using data, and that integrate data-driven decision making.

Big data and analytics

Big data poses both opportunities and challenges for businesses. In order to extract value from big data, it must be processed and analyzed in a timely manner, and the results need to be available in such a way as to be able to effect positive change or influence business decisions. The effectiveness also relies on an organization having the right combination of people, process and technology.

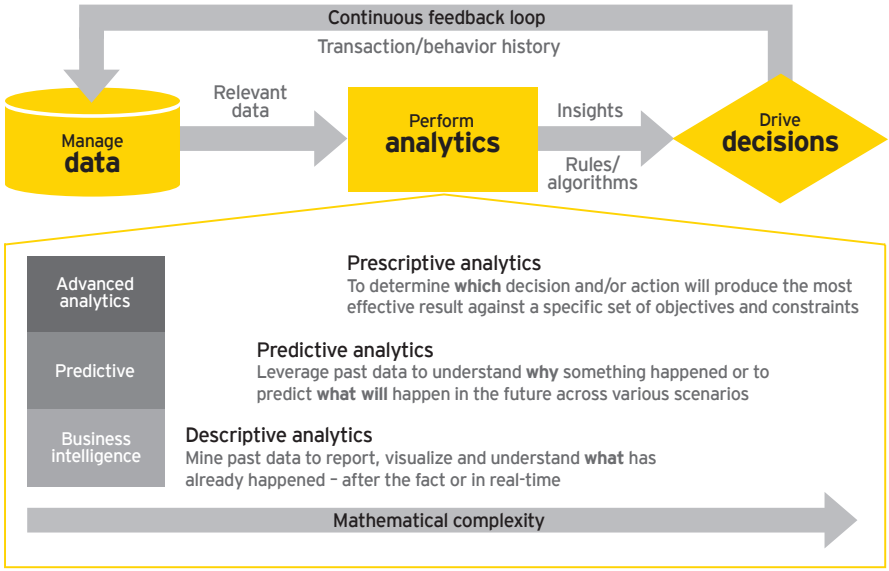
By pure definition, analytics is the discovery and communication of meaningful patterns in data – but for business, analytics should be viewed as the extensive use of data, statistical and quantitative analysis, using explanatory and predictive models to drive fact-based business management decisions and actions.

Analytics helps to optimize key processes, functions and roles. It can be leveraged to aggregate both internal and external data. It enables organizations to meet stakeholder reporting demands, manage massive data volumes, create market advantages, manage risk, improve controls and, ultimately, enhance organizational performance by turning information into intelligence.

EY analytics value chain

The goal is to use analytics to improve the **efficiency** and **effectiveness** of every **decision** and/or **action**.

1. Begin with leveraging leading tools and techniques to manage and extract relevant data from big data sources.
2. Applications of analytics can range from historical reporting, through to real-time decision support for organizations based on future predictions.
3. Use the insight generated by the analysis to drive change.



Analytics can identify innovative opportunities in key processes, functions and roles. It creates a catalyst for innovation and change — and by challenging the status quo, it can help to create new possibilities for the business and its customers. Sophisticated techniques can allow companies to discover root causes, analyze microsegments of their markets, transform processes and make accurate predictions about future events or customers' propensity to buy, churn or engage.

It is no longer enough for companies to simply understand current process or operations with a view on improving what already exists, when there is now the capacity to question if a process is relevant to the business, or whether there is a new way of solving a particular issue. The key driver for innovation within organizations is to constantly challenge existing practices rather than consistently accept the same.

Most organizations have complex and fragmented architecture landscapes that make the cohesive collation and dissemination of data difficult. New analytic solutions are playing an important role in enabling an effective Intelligent Enterprise (IE). An IE helps to create a single view across your organization by utilizing a combination of standard reporting and data visualization:

- Data from multiple source systems is cleansed, normalized and collated
- External feeds can be gathered from the latest research, best practice guidelines, benchmarks and other online repositories
- Use of enhanced visualization techniques, benchmarking indexes and dashboards can inform management and consumers via smartphones, laptops, tablets, etc., in-house or remotely

All companies need to start thinking about collecting and using relevant big data. Data-driven decisions can reduce inefficiency between the business, legal and IT, optimize existing information assets and address disconnects between different functions of an organization. However, it is worth noting that the best data and the most advanced analytical tools and techniques mean nothing if they are not being leveraged by people who are asking the right questions. Big data, emerging storage technology platforms and the latest analytical algorithms are enablers to business success — not a guarantee of it.

Big data can be a powerful way to identify opportunities, but when combined with traditional organizational information the volumes of data collected can be vast and traditional storage methods can be prohibitively expensive and do not necessarily scale effectively.

Organizational information is typically historical, incomplete and inaccurate. For a forward-looking perspective (using statistical and predictive modeling) it needs to be enriched with external information.

Big data drivers

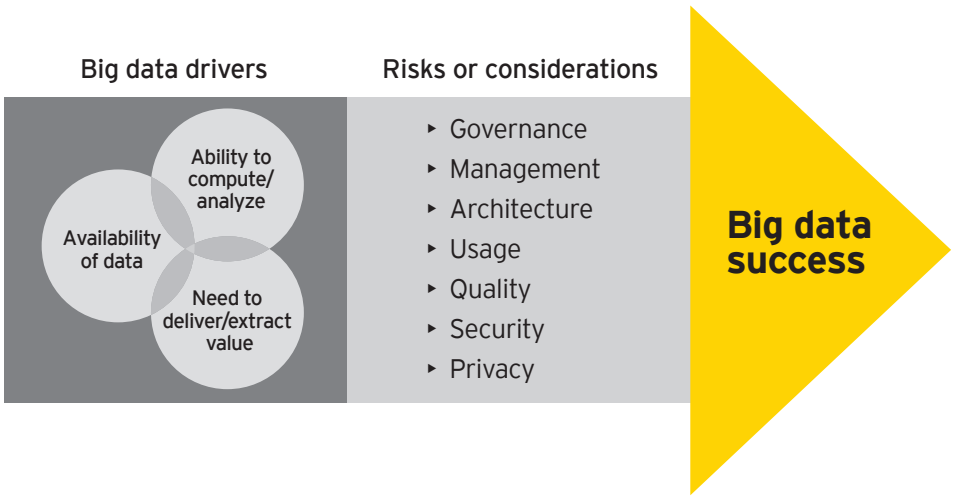
The benefits and risks of big data

While there is no doubt that the big data revolution has created substantial benefits to businesses and consumers alike, there are commensurate risks that go along with using big data.

The need to secure sensitive data, to protect private information and to manage data quality, exists whether data sets are big or small. However, the specific properties of big data (volume, variety, velocity, veracity) create new types of risks that necessitate a comprehensive strategy to enable a company to utilize big data while avoiding the pitfalls. This should be done in a prioritized fashion so that companies can start to realize the benefits of big data in step with managing the risks.

The following pages look at the possibilities and risks associated with big data and give examples of how big data is being leveraged to solve some of the complex issues businesses face today. We identify traditional and new risks and considerations for the seven key steps to success: governance, management, architecture, usage, quality, security and privacy.

A key success factor for companies is the availability of relevant information at the right time.



Governance

Good governance is vital to the success of Big data initiatives in any business; it encompasses consistent guidance, procedures and clear management decision-making. Organizations need to ensure standard and exhaustive data capture; they need not protect all the data, but they need to start sharing data with in-built protections with the right levels and functions of the organization.

Benefits

The MIT Center for Digital Business states "When it comes to big data, the 'right' governance model depends on the maturity level of the organization regarding data driven decisions." It obviously also highly depends on if big data is used to create new business or to drive more sales.

To unleash the power of big data, first of all data must be available and made fit for sharing. When it comes to, for example, medical data, respect for privacy and trust is inevitable.

Standardization in governance structures, with an integrated combination of technical, organizational and legal measures and safeguards, will help to increase trust. This is especially important when integrating governmental, institutional "open" and company data.

An example of this in action is seen with a European association that aims to build up a big data services platform for the health sector in their local region. It's a unique collaboration between health institutions, government, education and knowledge institutions and major IT service providers, addressing both the clinical and research sides of the health sector.

The core solution is comprised of a "vendor neutral hub" – a platform that works independently of vendors and data "owners" – where data can be captured, safely and durably stored, processed and distributed and finally, shared, if permitted by the data owner. This may be the case when patients are being treated by multidisciplinary teams or having treatment in various locations, or for research purposes using large sets of anonymous data.

The framework offers solutions for the immediate need to access data, to substantially lower the cost of data storage and, more importantly, to do more with the rapidly growing amount of unstructured data.

Risks and considerations

Traditional risks

- There is continued regulatory pressure on companies to meet a variety of policies and laws (e.g., Basel II, MiFID, SOX). Compliance governance is an expensive and complex problem to deal with, but failing to meet regulations can mean safety risks, hefty penalties, loss of reputation or even bankruptcy.
- In a global and continuously and rapidly changing legal and IT landscape it is not always clear exactly what legal and regulatory compliance entails (Who is responsible? Who is liable?), or how best to translate abstract rules from laws into organizational and technical measures within a company.
- Companies need to balance contradictive rules and regulations e.g., obligations based upon the US Patriot Act and the EU Data Protection Directive (and its many local implementations).

New risks

- Managers will need to learn to embrace the evidence-based decision-making process. Organizations have to redefine their understanding of "judgements" of the outcome of big data analytics.
- Data can be of great value, but companies have to consider ownership and privacy issues before using big data results. In the case of medical data, it is sometimes not clear who is the owner of the data, but using the data without the right legal foundation or consent of the patient may cause big problems.
- Big data may bring about intellectual property issues, e.g. copyright and database rights infringements. It will be a challenge to make sure that employees are not sharing inappropriate information, or too much data outside of the organization.

"Our top priorities are working with large data volumes and improving the efficiency of our testing. By using larger data sets we hope to do smarter internal audits, including more effective fraud detection."

Head of Internal Audit, Australia

Management

Integrating and moving data across the organization is traditionally constrained by data storage platforms such as relational databases or batch files with limited ability to process very large volumes of data, data with complex structure or without structure at all, or data generated or received at very high speeds.

Organizations need to start managing data through different sources, and integrating its usefulness via a range of technologies in the market.

Benefits

Big data overcomes traditional restraints in a cost-effective manner and opens opportunities to ingest, store and process data from new sources such as external social media data, market data, communications, interaction with customers via digital channels, etc.

By some estimates, more than 80% of the data within organizations is unstructured and unfit for traditional processing. Using big data will enable the processing of this unstructured data and increased system intelligence which can be used to improve performance in sales, increase understanding of customer needs, reinforce the internal risk management function, support marketing initiatives and enhance fraud monitoring.

Big data capability allows organizations to integrate multiple data sources with relatively low effort in a short timeframe. Combined with a lower cost of storage per gigabyte, this enables organizations to build, for example, a federated view of customers by shifting customer data from various separate business departments into a single infrastructure, and then to run consolidated analytics and reporting on it.

Big data technologies release organizations from the traditional accuracy vs. cost challenge by enabling them to store data at the lowest level of detail, keeping all data history under reasonable costs and with less effort.

Risks and considerations

Traditional risks

- There is a long implementation cycle for data warehousing and reporting solutions.
- Challenges over unifying data definitions are made even more complex across multiple business lines.
- Modeling, storage and processing challenges arise from the growing volumes of data with dynamic structures.

New risks

- Simplified access to diverse sources of data and easy-to-ingest large amounts of information may result in increasing amount of "noise" in data and decrease in the overall level of data quality.
- Many new technology market players don't have mature enterprise-ready capabilities around implementation, support, training, etc.
- New big data methods, architecture and volume variety impose additional risks of lack of control and governance over data, and this requires additional organizational focus. Under the context of the complex data landscape, it is especially important to establish and maintain data lineage.
- Organizations may struggle with finding the right skills and building internal capabilities for handling big data as most of the technologies and methods are relatively new, and market resources are in short supply.

Example engagement: management

In response to the recent Dodd-Frank regulations, requiring financial organizations to report all pre-trade communications data across the organization related to a trade under query within short notice, EY has developed a solution that allows organizations to harvest, index and link unstructured information related to internal communications data to related trades.

The solution is able to process large variety of unstructured pre-trade communications data sources (emails, instant messages, phone calls, etc.) and applies a number of matching rules and fuzzy logic to match this data to financial transactions the communications relate to: this enables the organization to fetch the full history of communication events related to a specific transaction almost instantly.

While addressing the regulatory requirement, the solution also provides essential capabilities for rogue trader analytics as it effectively supplements traditional analytics models that leverage trade economics data with unstructured pre-trade communications data feed that significantly broadens the context and precision of analysis.

Architecture

With big data, it has become possible to build an architecture which can integrate massive volumes of data in various formats and provide real-time analytics aimed at a consolidated customer view, or improved fraud detection and other similar business goals.

Data architecture should be prepared to break down internal silos, enabling the sharing of key data sets across the organization and to ensure that learnings are being captured and relayed across to the right set of people in the organization in a timely and accurate manner.

Benefits

Big data has brought a new paradigm to data architecture. In the past, data systems were built with a predetermined set of data requirements. In the Big Data world, data storage platforms are not restricted to a predefined rigid data model, and data systems are capable of handling all kinds of structured and unstructured data.

Integration of unstructured data in particular can lead to improved analytics and reporting. For example, a business goal of having a consolidated view of the customer profile across business functions and geographies is important for various reasons:

- To make the business decision making process more intelligent
- To enhance the monitoring of customer profiles for "red flags" (issues of concern or opportunity)
- To enable the company to offer more relevant services to their customers tailored to their specific needs

Traditionally, organizations struggled to achieve this goal because their customer data was lying in multiple systems and different file formats (PDFs, Word and Excel documents, charts, images, scans, videos, etc.): technology was seen as a limiting factor to integrate this scattered and massive data and meet the goal. Big data brings a solution to this by offering capabilities to integrate and analyze data coming from large variety of systems across the organization in an efficient and flexible manner.

Real-time fraud monitoring is a classic big data challenge, demanding the integration of large amounts of diverse, structured and unstructured high-velocity data that needs to be analyzed in near-real time to realize the benefits. A global payments technology company recently stated that it has made an improvement of 130% in identifying fraud for debit transactions and 175% improvement in cases of credit card transactions by using big data technologies authorization model.

Big data also offers additional capabilities such as deploying data storage/processing power over a grid of commodity hardware, with unconstrained scalability and flexibility to adapt to constantly changing data landscape.

Risks and considerations

Traditional risks

- Increasing the volume of data puts a strain on infrastructure, resulting in slow processing, storage problems and back up requirements.
- The inability to work with unstructured data reduces the quality of analytics and reporting.
- Numerous data silos create the risk of poor data integrity, inconsistency and high implementation and maintenance budgets.

New risks

- More is not always better. More data can lead to an increased number of data quality issues, and confusion and lack of consistency in business decision making – especially when conflicting information is present.
- Integrated data architecture increases the challenges of data linkages and matching algorithms to distinguish items of relevance from piles of data.
- Increased complexity of architectural landscape and the growing amount of data bring new challenges around data governance and data privacy.
- Lack of capabilities, both within organizations and externally, make it hard to keep up with rapidly evolving hardware/software technology and implementation methods.

Example engagement: architecture

A top 10 global insurance company invested $300 million in a futuristic big data solution aimed at providing a single customer view across their enterprise; providing a 360 degree view of their customer portfolio and consolidates customer interactions.

Before the big data solution, the organization struggled to create a single view of its customers across product and business systems because of growing volumes, incompatible systems, inconsistency of data and inability to process unstructured data. They are now able to generate a single customer view which can be used in a consistent manner across the enterprise, leading to effective sales recommendations, customized offerings and operational improvements.

Usage

The convergence of data availability and processing power is helping to unlock the potential of big data for most sectors and industries. The results of big data can beneficial to a wide range of stakeholders across the organization — executive management and boards, business operations and risk professionals, including legal, internal audit, finance and compliance; as well as customer-facing departments like sales and marketing.

The key challenge is having the ability to interpret the huge amount of data that can be collated from various sources.

Benefits

The weather used to be unforeseeable and ungovernable. Robust weather forecasts models usually require hundreds of thousands of atmospheric variables that are constantly changing. With big data, some technology companies have emerged with the ability to provide historical weather data and better forecasting of extreme weather events. Based on billions of calculations and data points over the past several decades, big data now makes it possible to improve weather predictions up to a month in advance.

With the advent of low-cost cloud computing environments and open data movements, various big data weather forecasting ventures have arisen in recent years. Some of those new start-ups provide their services to corporate users (e.g., large-scale farmers, logistic companies) and some to retail customers directly.

Accurate weather data is beneficial for many organizations; for example, some companies have been using weather information to improve their business activities ranging from supply chain planning to advertising.

Supply chain management goes beyond just stocking more shovels ahead of a snowstorm. Retailers can now improve inventory management by leveraging new big data insights showing, for example, that after unusually cool weather, beer sales will decrease in some cities while increasing in others. And, by combining real-time detailed analysis of current and historical weather data with personal data such as location, demographics and purchase history, retailers are able to further refine and target their advertisements; i.e., consumer purchase patterns will change if today is the first warm day after a week of cold temperatures.

Risks and considerations

Traditional risks

- A key challenge is to know the right business questions to ask.
- There are misunderstandings over what data is needed to make strategic or operational decisions.
- Many organizations do not have the ability to analyze data timely enough to take advantage of new insights.

New risks

- Not considering information from outside the organization (e.g., weather) that is relevant to answering bigger question is an ongoing concern.
- There is a shortage of qualified "data scientists" globally for the near to mid-term.
- Organizations can get overloaded and overwhelmed by trying to handle too much data.
- The challenge of getting the right information to the right person at the right time is expanded due to the sheer size of big data.
- The costs associated with managing and monitoring the quality, credibility and integrity of big data can be prohibitive.
- There is a necessity to temper the expectation that big data will solve everything.

Big data brings value to companies in diverse and unexpected ways.

"Companies that are moving ahead are doing an incredible amount of analysis around selection and recruitment processes and staff turnover to figure out if they can identify what is happening with their workforce. The focus on data is a real winner for HR if handled correctly."

Edward Lawler, Director of the Center for Effective Organizations and Distinguished Profession of Business at the University of Southern California

Quality

The quality of data sets and the inference drawn from such data sets are increasingly becoming more critical and organizations need to build quality and monitoring functions and parameters for big data. For example, correcting a data error can be much more costly than getting the data right the first time — and getting the data wrong can be catastrophic and much more costly to the organization if not corrected.

Benefits

For many years the health care ecosystem has embraced big data. With the ability to capture every patient touch point, the amount of data within the health care ecosystem has exploded. The evolution of new data sources and the ability to mash that data with existing data sources is evolving — big data is creating the possibility of new positive patient outcomes.

Some of these new data sources include the integration of disease registries, tissue registries and genomic information, and then aligning them with meaningful use clinical standards. It is defining key care treatment approaches based on new genetic insights and clinical protocol matching algorithms, and defining focused patient care treatment insights earlier within the care delivery process.

The value from these new big data insights will be priceless for the patient. The quality of the data will also have a direct effect on driving new key health care insights in creating high-quality outcomes while effectively managing costs.

Risks and considerations

Traditional risks

- A key concern is building and maintaining golden sources of data and determining which data sources should be golden.
- There is an ongoing risk in maintaining "fit for purpose" status (i.e., completeness, validity, consistency, timeliness and accuracy) of enterprise data as it moves from the source to reporting.
- Understanding the data domains and the level of data quality required for each data domain creates risk.
- Well-established information governance models do not exist to manage data quality.

New risks

- Companies may not have the knowledge or competencies to deal with big data.
- The need to interpret and assess unstructured data can be a challenge, and the quality of the unstructured data is often unproven.
- Structured and unstructured data may not be integrated cohesively.
- Existing information governance models will not be aligned to manage data quality for the newly acquired data.

Common FDA investigations and compliance monitoring include:

- **Payment stream, accounts payable analyisis**
  - Altered invoices, duplicate or fake invoices, inflated prices, suspicious payment descriptions, requestor/approver conflicts
- **Vendor master/employee master analyisis and comparisons**
  - Fictitious vendors, vendor risk ranking, background due diligence, conflicts of interest
- **Employee expenses, travel and entertainment**
  - Over limits, unusual or inappropriate expenses, miscellaneous/sundry expenses, split or duplicate expenses
- **Payroll**
  - Ghost employees, falsified wages, commission schemes
- **Financial misstatement**
  - Ficticious revenues, bill-and-hold schemes, concealed liabilities, imporper discosures, overstated assets
- **Bribery and correuption**
  - Bid rigging, conflicts of interest, contract compliance, kickbacks, illegal gratuities
- **Capital projects**
  - Contract noncompliance, project abuses and overcharges

Security

Security is a major concern with big data. To make more sense from the big data, organizations would need to start integrating parts of their sensitive data into the bigger data. To do this, companies would need to start establishing security policies which are self-configurable: these policies must leverage existing trust relationships, and promote data and resource sharing within the organizations, while ensuring that data analytics are optimized and not limited because of such policies.

Benefits

With the rapid adoption of smartphone and wearable technology, organizations are increasingly taking advantage of the new medium (e.g., developing their own apps, partnering with a third-party app, or purchasing advertising space within a third-party app) in which to reach out to customers to increase their revenue.

For example, health and fitness organizations have developed mobile apps that utilize the GPS function of a smartphone and motion/accelerometer to determine where you are, plot your route on a map, show your running speed, etc. Some organizations have also developed software coupled with wearable technology that monitors your sleeping patterns.

Other location-based apps are selling GPS-based advertising within their application. This, for example, can enable retail organizations to target product offers, and offer discounts to consumers within the given radius of the retail store.

In order for this type of data to be useful, it needs to be linked to an entity. The linkage may be to a user account that identifies a person directly, or a device that may be traced to a person. The collection and aggregation of this type of information increasingly exposes an individuals' information. Maintaining security over this data is integral to building consumer trust and fostering continued usage.

Risks and considerations

Traditional risks

- Historically, data such as your fitness regime, sleep patterns and habits, etc. may have been collected via anonymous surveys; therefore, protecting that data was restricted to protecting an organization's market research.
- Centralized storage meant that data could be protected via defense in depth techniques such as firewalls, access management, anti-virus software; however, a breach still can result in loss of massive amounts of data.
- Even with meticulous internal access policies and procedures, it is difficult to defend against rogue users who intentionally lead information.

New risks

- Diverse sources of data results in distributed storage and management, compounding security vulnerabilities.
- Cloud computing puts more data in motion and causes additional security complexities.
- Lost data results in more direct impact to the end consumer. This, coupled with the behavioral data collected, can lead to more sophisticated attacks, e.g., social engineering attempts on targeted consumers.
- Increasing global regulations raise the stakes around security as the cost of dealing with data breaches continues to grow.

"Security intelligence is the key to the future. We need big data techniques to find the bad guys."

Financial services organization

Privacy

Organizations have traditionally used various methods of de-identification (anonymization, encryption, pseudonymization, key-coding, data sharding, etc.) to distance data from real identities and allow analysis to proceed while at the same time containing privacy concerns.

The increased use of big data challenges the traditional frameworks for protecting the privacy of personal information, forcing companies to audit the implementation of their privacy policies to ensure that privacy is being properly maintained.

Benefits

Automobiles have evolved from purely mechanical machines to sophisticated mobile platforms that host a variety of computer, sensors and GPS and communications devices.

The use of sensor and location data in the automotive industry has increased tremendously in the past few years. Services such as subscription-based communication and navigation systems provide safety and convenience to users and new revenue streams for companies. The ability to view car status such as fuel level and tire pressure on a smartphone, track a stolen vehicle or summon assistance in case of accidents are all enabled by real-time collection and monitoring of location and sensor data.

Using crowd-sourced real-time location information, mapping apps and GPS units provide more accurate traffic and drive-time information, and by analyzing the commute patterns of tens of millions of drivers, actual drive times to and from work can now be factored into home purchase decisions.

Even telematics information such as braking and cornering force is being combined with data, such as when you drive (e.g., after midnight) and how often the speed limit is exceeded, in order to determine personalized auto insurance premiums.

Risks and considerations

Traditional risks

- Privacy considerations around personal information have always existed, but companies have traditionally dealt with them on a limited or single location basis.
- Sharing information with third parties (e.g., a subcontractor that services client requests, or data being sold to other companies) creates the need for increased privacy considerations.

New risks

- New technology enables the ability to gather and analyze data from multiple locations and brings the challenge of meeting different privacy regulations around the world: similar data, such as telematics, may need to be treated differently depending on where it is generated or captured.
- Sensor and geolocation data may be used to identify an individual, even if no name is attached to the data, therefore increasing the need to define "personal information" that needs to be protected.
- The ability to collect new sources of information, like car-based sensor data, increases the need to evaluate opt-in/opt-out procedures by consumers.

Example engagement: privacy

As part of its plan to develop a consistent global privacy program, a leading manufacturer commissioned EY to conduct a study around global privacy definitions, regulations and incidents. The research was centered on industry-specific examples of data collection and usage, including precedents where companies had been found to be misusing personal information.

The project highlighted some of the fundamental differences in the approaches taken by US, European, Asian and other countries toward recognizing and protecting personal information, with more strict definitions and protection requirements in Europe. This information formed the foundation for a re-alignment of the company's global privacy policies, and EY continues to assist with insights around regulations and the life cycles of data.

The bigness of big data is growing faster than businesses realize. Today, we generate more data in 10 minutes than all of humanity has ever created through to the year 2003.

Big data has big potential

Organizational information is typically historical, incomplete and inaccurate. For a forward-looking perspective (using statistical and predictive modeling) it needs to be enriched with external information (big data). However, traditional systems and approaches are slow and inflexible and cannot handle the new volume and complexity of big data.

A key success factor for companies is the availability of relevant information at the right time. Businesses need to know *what* decisions should be made, *when* to take action and *how* these decisions will impact on financial results and operational performance. Demand for this type of insight fuels the growth of big data to enable them to make better, smarter, real-time, data-driven decisions that will change the way they handle their operations and compete in the marketplace.

The field of big data is evolving rapidly and organizations cannot ignore the benefits. At the same time, they must look carefully at the way they use the data to ensure they are controlling the new opportunities while managing the new risks.

Part of the challenge for organizations in successfully executing a big data strategy is to develop sound fundamentals that are flexible enough to address the organization's data requirements of today and tomorrow.

Lagging	Industry trends	Leading
	Data	
Functional silos ←	Users	→ Spans enterprise
Internal ←	Sources	→ Internal/external
Multiple ←	Technology/repository	→ One trusted
	Analytics	
Old patterns ←	Purpose	→ Exploratory
Several ←	Variables	→ Many
Historic trends ←	Discovery	→ Predictive
	Decisions	
"Gut feel" ←	Based on	→ Data
Tenure ←	Business case development	→ Data
Perception ←	Road map prioritization	→ Data

Next moves

EY has helped many global organizations to successfully deploy the right people, processes and technologies around data issues and solutions.

To become more competitive and more efficient, companies need to look at the broader set of related risks, incorporate more data sources, use better tools to allow them to move to real-time or near-real-time analysis and increase data volumes. We suggest that organizations need to consider the following key questions when assessing their readiness to truly start benefiting from big data.

Key questions		Yes	No
Governance	Given the ubiquitous nature of big data, does your data governance framework acknowledge the evolving definitions of data owners and consumers?		
	Does your current governance address the risks related to the life cycle of big data?		
Management	Do you have the right skills and internal capabilities to deal with the big data technologies and methods which are relatively new?		
	Do you have sufficient control over the big data volumes, variety, velocity and veracity, which may impose additional risks?		
Architecture	Does your IT infrastructure support your big data strategy?		
	Can you flexibly scale processing and storage to meet the demands of big data processing?		
Usage	Do you have the right talent to be able to process, model and interpret big data results?		
	Is your workforce ready to shift to the new paradigm of data-driven decisions?		
Quality	Are your existing methods sufficient to deal with the unstructured data?		
	What level of data quality is required to meet your big data goal?		
Security	Is your security infrastructure robust enough to deal with the increasing demands of protecting a growing stockpile of data, while flexible enough to not become bottlenecked by the innovation?		
Privacy	Have you defined who owns big data information, and whether there is actual or implied consent to use the same?		
	Do you understand that how the big data is stored and how it is used can also create significant privacy issues?		

If one of these questions has been answered with "No," it is time for you to take action.

Want to learn more?

Insights on governance, risk and compliance is an ongoing series of thought leadership reports focused on IT and other business risks and the many related challenges and opportunities. These timely and topical publications are designed to help you understand the issues and provide you with valuable insights about our perspective.

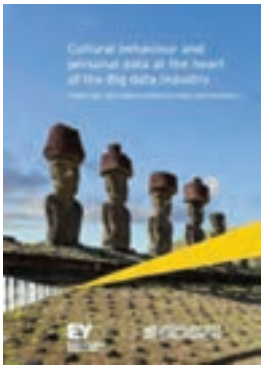
Please visit our Insights on governance, risk and compliance series at www.ey.com/GRCinsights

Under cyber attack: EY's Global Information Security Survey 2013.
www.ey.com/giss2013

Privacy trends 2014: Privacy protection in the age of technology.
www.ey.com/privacy2014

Big risks require big data thinking: EY's Global Forensic Data Analytics Study 2014.
www.ey.com/FDASurvey

Big Data and Enterprise Mobility: Growing relevance of emerging technology themes – The India perspective.
www.ey.com/bdem



Cultural behaviour and personal data at the heart of the Big data industry: Finding the right balance between privacy and innovation.
www.ey.com/bigdataprivacy

Ready for takeoff?: Overcoming the practical and legal difficulties in identifying and realizing the value of data.
www.ey.com/readyfortakeoff

"For us, the biggest challenge is the integration and acceptance of analytical tools within our organization. From an integration perspective, it's about how we integrate those tools with our existing technology; from an acceptance perspective, it's about ensuring that the benefits of prevention are fully understood within the executive management community."

Head of Internal Audit, Australia

At EY, we have an integrated perspective on all aspects of organizational risk. We are the market leaders in internal audit and financial risk and controls; and we continue to expand our capabilities in other areas of risk, including governance, risk and compliance as well as enterprise risk management.

We innovate in areas such as risk consulting, risk analytics and risk technologies to stay ahead of our competition. We draw on in-depth industry-leading technical and IT-related risk management knowledge to deliver IT controls services focused on the design, implementation and rationalization of controls that potentially reduce the risks in our clients' applications, infrastructure and data. Information security is a key area of focus where EY is an acknowledged leader in the current landscape of mobile technology, social media and cloud computing.

About EY
EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit ey.com.

EYG no. AU2334
ED none

In line with EY's commitment to minimize its impact on the environment, this document has been printed on paper with a high recycled content.

ey.com/GRCinsights

About EY's Advisory Services

Improving business performance while managing risk is an increasingly complex business challenge. Whether your focus is on broad business transformation or more specifically on achieving growth, optimizing or protecting your business, having the right advisors on your side can make all the difference.

Our 30,000 advisory professionals form one of the broadest global advisory networks of any professional organization, delivering seasoned multidisciplinary teams that work with our clients to deliver a powerful and exceptional client service. We use proven, integrated methodologies to help you solve your most challenging business problems, deliver a strong performance in complex market conditions and build sustainable stakeholder confidence for the longer term. We understand that you need services that are adapted to your industry issues, so we bring our broad sector experience and deep subject matter knowledge to bear in a proactive and objective way. Above all, we are committed to measuring the gains and identifying where your strategy and change initiatives are delivering the value your business needs.

To find out more about our Risk Advisory services could help your organization, speak to your local EY professional or a member of our global team, or view: ey.com/advisory

The leaders of our Risk practice are:

Global Risk Leader		
Paul van Kessel	+31 88 40 71271	paul.van.kessel@nl.ey.com
Area Risk Leaders		
Americas		
Jay Layman	+1 312 879 5071	jay.layman@ey.com
EMEIA		
Jonathan Blackmore	+44 20 795 11616	jblackmore@uk.ey.com
Asia-Pacific		
Iain Burnet	+61 8 9429 2486	iain.burnet@au.ey.com
Japan		
Yoshihiro Azuma	+81 3 3503 1100	azuma-yshhr@shinnihon.or.jp