Big Data in Europe: Evolution AND Revolution

EITO

single user license

Imprint

European Information Technology Observatory 2013/14

Publisher

Bitkom Research GmbH
Albrechtstraße 10, 10117 Berlin, Germany
Register of companies: Amtsgericht Charlottenburg
HRB 87370 B
Managing Director: Dr. Axel Pols
Phone: +49.30.27576-560
Fax: +49.30.27576-51560
Internet: www.bitkom-research.de

EITO Sponsors

Deutsche Telekom AG
KPMG AG
Telecom Italia S.p.A.

With the support of:

European Commission, DG Connect, DG Enterprise and Industry
OECD, Directorate for Sience, Technology and Industry

Research Partners

The EITO Report 2013/14 has been jointly prepared by IDC and the EITO Task Force on the basis of the information available as of October 2013. IDC provided data and analysis for IT and telecommunications markets.

EITO Task Force

Enrico Barella (Telecom Italia)
Ronald Chodasz (FEEI)
Antonio Cimorra (AMETIC)
Elizabeth Cobbson (Intellect)
Diane Dufoix (AFDEL)
Dana Eleftheriadou (EU Commission)
Andreas Fier (Deutsche Telekom AG)
Petar Indovski (MASIT)
Pierre Muckly (SWICO)
Michael Müllneritsch (FEEI)
Carmelo Javier Muñoz Ruiz (red.es)
Carlos Perez-Maestro (European Commission)
Axel Pols (BITKOM e.V.), Chairman
Taylor Reynolds (OECD)
Loïc Rivière (AFDEL)
Tiziano Rognone (ANITEC)
Yannis Sirros (SEPE)
Patrick Slaets (AGORIA)
Myriam Vassiliadou (SEPE)
Bruno Wallraf (KPMG)
Verena Weber (OECD)

EITO Project Management

Franz Grimm
Katja Hampe

Layout and Graphics

Astrid Scheibe

Editorial Support

Bartlomiej Kowalik

The EITO project is sponsored by following companies and organizations (Gold sponsors):

Deutsche Telekom AG
www.telekom.de

Telecom Italia S.p.A.
www.telecomitalia.it

KPMG Deutsche Treuhand-Gesellschaft
Aktiengesellschaft, Wirtschaftsprüfungsgesellschaft
www.kpmg.com

The EITO reports are prepared and distributed in cooperation with our partner organizations across Europe

Afdel (France)
L'Association Française Des Éditeurs De Logiciels
www.afdel.fr

AGORIA (Belgium)
Belgium Federation for the technology industry
www.agoria.be

AMETIC (Spain)
Asociación multisectorial de empresas de la electrónica, las tecnologías de la información y la tecomunicación, de las telecomunicaciones y de los contenidos digitales
www.ametic.es

ANITEC (Italy)
Associazione Nazionale Industrie Informatica, Telecomunicazioni ed Elettronica di Consumo
www.associazioneanitec.it

BITKOM (Germany)
Bundesverband Informationswirtschaft, Telekommunikation und neue Medien
www.bitkom.org

FEEI (Austria)
Fachverband der Elektro- und Elektronikindustrie
www.feei.at

intellect (UK)
Information Technology Telecommunications & Electronics Association
www.intellectuk.org

MASIT (Macedonia)
Macedonian ICT chamber of commerce
www.masit.org.mk

SEPE (Greece)
Federation of Hellenic Information Technology & Communications Enterprises
www.sepe.gr

SWICO (Switzerland)
Schweizerischer Wirtschaftsverband der Anbieter von Informations-, Kommunikations- und Organisationstechnik
www.swico.ch

red.es (Spain)
Corporate entity attached to the Spanish Minstry of Industry, Energy and Tourisms
www.red.es

Table of content

List of figures

Note: This Report has been prepared by IDC in cooperation with the EITO Task Force.

Executive Summary

- Big Data is becoming mainstream in North America, but Europe has lagged behind until now. There are three main reasons for this: European organizations and hence their datasets are generally smaller, so on average organizations feel less pain from the lack of information; reluctance to hire new people in an area suffering a dramatic skills shortage; and the economic climate which, although its impact on ICT spend across the board has been slight, it has made organizations more cautious when looking at a new area like Big Data.

- However, activities are starting to pick up, driven by the increase in information and awareness of best practices coming from the North American market, the rapidly evolving enterprise versions/standards around open source frameworks like Hadoop and Cassandra beyond the very early stages, and adoption in highly advanced segments such as technology startups and video game developers. This adoption will be driven further in 2013 as more enterprise Hadoop vendors start to enter the European market.

- The market hype around Big Data is substantial, and continues to grow. Big Data is in one sense an evolution of what business analytics has been doing (under various names) for over two decades. However, from a technology perspective, Big Data comprises a set of genuinely new technologies (examples are Hadoop, highly scalable databases, advanced data visualization tools and high-performance search engines) and a convergence of more mature technologies (examples are event-driven processing, business intelligence and data mining). One thing that is certain from two decades of experience with business analytics: to fully embrace Big Data, organizations need to be dedicated and determined to embrace a more information-led culture.

1 Situation Overview

1.1 The Information Explosion

- For five years starting in 2007, IDC conducted an annual study, sponsored by EMC, to size the digital universe, which is the amount of data on the planet. The December 2012 version of the study generated the following insights:

  - From 2005 to 2020, the digital universe will grow by a factor of 300, from 130 exabytes to 40,000 exabytes, or 40 trillion gigabytes (more than 5,200 gigabytes for every man, woman, and child in 2020). From now until 2020, the digital universe will about double every two years.

  - The investment in spending on IT hardware, software, services, telecommunications and staff that could be considered the "infrastructure" of the digital universe and telecommunications will grow by 40% between 2012 and 2020. As a result, the investment per gigabyte (GB) during that same period will drop from $2.00 to $0.20. Of course, investment in targeted areas like storage management, security, big data, and cloud computing will grow considerably faster.

  - A majority of the information in the digital universe, 68% in 2012, is created and consumed by consumers — watching digital TV, interacting with social media, sending camera phone images and videos between devices and around the Internet, and so on. Yet enterprises have liability or responsibility for nearly 80% of the information in the digital universe. They deal with issues of copyright, privacy, and compliance with regulations even when the data zipping through their networks and server farms is created and consumed by consumers.

  - Only a tiny fraction of the digital universe has been explored for analytic value. IDC estimates that by 2020, as much as 33% of the digital universe will contain information that might be valuable if analyzed.

- By 2020, nearly 40% of the information in the digital universe will be "touched" by cloud computing providers — meaning that a byte will be stored or processed in a cloud somewhere in its journey from originator to disposal.

1.2 What is Big Data?

- IDC defines Big Data technologies as a new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis.

- This definition encompasses hardware and software that integrates, organizes, manages, analyzes, and presents data that is characterized by the "four Vs":

  - Volume
  - Variety
  - Velocity
  - Value

- The importance of each of these attributes varies depending on the specifics of each industry or even organization. While the first three attributes of Big Data are defined in terms of technical characteristics, the fourth attribute, value, is defined by the perceived value of the data and the technology to any given organization.

Volume: Size Is Not the Only Thing That Matters

- While the word „Big" in Big Data alludes to massive volumes of data, as noted previously, users must understand this as a relative term. Some industries and organizations are likely to have mere gigabytes or terabytes of data as opposed to the petabytes or exabytes of data for some of the social networking organizations. Nevertheless, these seemingly smaller applications may still require the intense and complex information ion processing and analysis that characterize Big Data applications.

- The financial services industry demonstrates this variability. When engaging in certain Big Data activities, there may be millions or billions of records to consider, but each record may only be several bytes long (such as stock ticker information). Conversely, email archives may accumulate several petabytes of data containing valuable customer suggestions or complaints, records of projects, legal records, contracts, and proposals. The email archive usually contains the best record of pending and current business, but it needs to be sorted and mined to find out what it contains. Another good example is product design and manufacturing, where automotive and aerospace companies, for example, may evaluate hundreds or thousands of virtual prototypes to home in on the best vehicle design. The new large-scale scientific experiments that generate petabytes of mixed data a day as input into a complex simulation model are another example.

Variety: The Combination of Data Sources and Formats Is What Matters

- Variety in Big Data is a critical attribute. The combination of data from a variety of data sources and in a variety of formats is a key criterion in determining whether an application can be considered as Big Data.

- Big Data applications typically combine data from a variety of data sources (typically both internal and external to an organization) and of different types (structured, semi structured, and unstructured). This is an important facet of Big Data for both technical and potential impact reasons. Combining types of information is a complex technical challenge: What is the relative importance of a tweet versus a customer record? How do you combine a large number of changing patient records with published medical research and genomic data to find the best treatment for a particular patient?

- An example of this may be the mash up of internal operational data from the ERP system with semi structured data from Web log files that identifies customers' online behavior, with sentiment analysis of unstructured text from customer comments.

- Another example is advanced weather/climate modeling that draws on 100 years of weather data with new physical models of ocean water behaviors and CO level changes, mixing in satellite data feeds to create a real-time simulation.

Velocity: Speed at Which Information Arrives and Is Analyzed and Delivered

- The velocity of data moving through the systems of an organization varies from batch integration and loading of data at predetermined intervals to real-time streaming of data. The former can be seen in traditional data warehousing and is also the primary method of processing data using Hadoop today. The latter is the domain of technologies such as complex event processing (CEP), rules engines, text analytics and search, inferencing, machine learning, and event-based architectures in general.

- The key to evaluating velocity requirements of Big Data is to understand the business/organizational processes and requirements of end users. For example, hedge funds need to grab emerging investment trends immediately — seconds can make a difference. Similarly, real-time face recognition is a requirement for airport screening of travelers to catch criminals as they enter airports. However, Web search engines — the source of MapReduce and Hadoop — must process and mine billions of queries to determine the accuracy of their algorithms or ad matching but don't need to perform that analysis in real time. In other words, the right information at the right time with the right degree of accuracy is what's needed.

- The technology infrastructure for each use case differs. There is an old adage in the infrastructure community — you can solve any problem if you throw enough hardware at it. And, when one considers the large supercomputers that have been built, or the massive clusters that have been created to address specific problems, that adage holds true.

- Increasingly however in today's world, the need for specialized hardware is not always necessary to meet high-performance demands. The combination of high-availability clustering, scale-out file systems, multi-CPU, and multicore processors means that the performance that can be delivered today levering common off-the-shelf (COTS) components is likely to be sufficient. Socialytics applications are often delivered via cloud, making consideration of the hardware unnecessary. This is important, particularly as discussed in the next section.

Value: Capital, Operational, and Business Benefits All Matter

- In the context of Big Data, value refers to both the cost of technology and the value derived from the use of Big Data. The cost variable is important because it is a key defining factor of what's new with big data. Large data warehouses in financial services, telecommunications, retail, R&D, and government organizations have existed for years. Real-time data management in trading, weather monitoring, or fraud detection applications has existed for years. Unstructured content analysis in the form of text mining has existed for years. High-performance computing systems for scientific research have existed for years.

- What is different now is that systems that were previously affordable and available only to government agencies or to a few of the largest companies in select industries are now available to the broader market. A combination of open source software and decreasing hardware prices has made these technologies more affordable.

- Value also refers to the benefits derived from Big Data projects. These benefits can be broadly classified as:

  - **Capital cost reduction.** A reduction in software, hardware, and other infrastructure costs

  - **Operational efficiency.** A reduction in labor costs due to more efficient methods for data integration, management, analysis, and delivery

- **Business process enhancements.** An increase in revenue or profit due to new or better ways of conducting business, including improvements to commercial transactions, sustainable management of communities, and appropriate distribution of social, healthcare, and educational services

IDC's Big Data Architecture

- The technologies to store, organize, analyze and deploy Big Data are divided by IDC into four types – infrastructure, data organization & management, analytics & discovery, and applications.

1.3 A Bird's Eye View of Big Data Adoption in Europe

Big Data is Picking up in Europe after a Slow Start

- Big Data is becoming more mainstream in North America, but Europe has lagged behind until now. There are three main reasons for this: European organizations and hence their datasets are generally smaller, so on average organizations feel less pain from the lack of information; reluctance to hire new people in an area suffering a dramatic skills shortage; and the economic climate which, although its impact on ICT spend across the board has been slight, it has made organizations more cautious when looking at a new area like Big Data.

- However, activities are starting to pick up, driven by the increase in information and awareness of best practices coming from the North American market, the rapidly evolving enterprise versions/standards around open source frameworks like Hadoop and Cassandra beyond the very early stages, and adoption in highly advanced segments with a high presence in Europe, such as video and online game development. This adoption will be driven further in 2013 as more vendors start to enter the European market (MapR) or ramp up (Cloudera opened its first European office in London in December 2012, and Hortonworks opened in London in March 2013 having been present in Continental Europe for a few years).

1.4 Big Data in Europe in 2013

- This study refers to the results of IDC's European Software Survey 2013. The survey was conducted in February 2013 and used computer-aided telephone interviews (CATI) to survey 700 end users across seven countries/regions in Western Europe.

- The respondent base consists of the following segments:

  - Country/region: France, Germany, Italy, the Netherlands, the Nordic region, Spain, and the U.K.

  - Industry: Financial services, manufacturing, retail and wholesale, public sector, and telecoms and media. An "other service industries" sector make up the total to 700 respondents.

  - Size class: 50–99 employees, 100–499 employees, 500–999 employees, 1,000–2,499 employees, and 2,500+ employees.

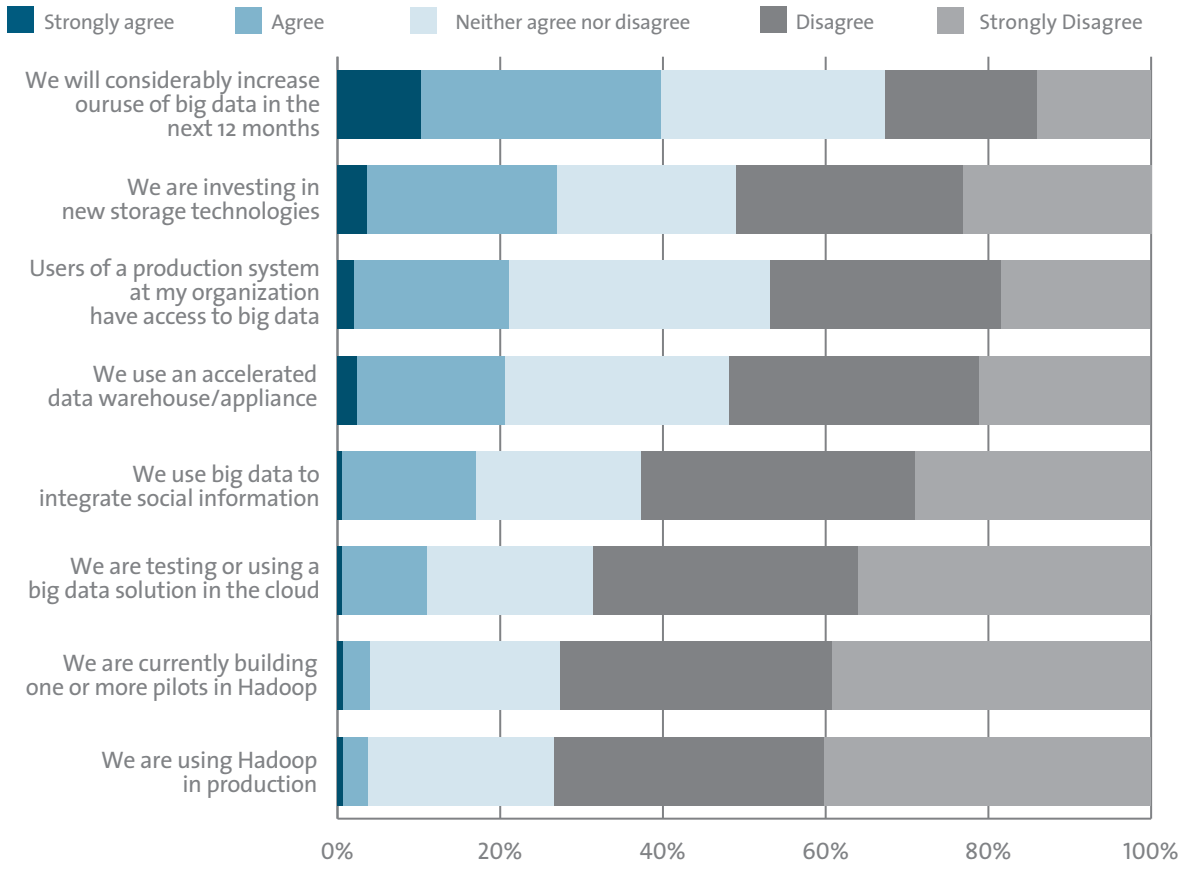
- For the Big Data part of the survey, respondents were asked to state their agreement with a series of Big Data-related statements. In the figures and tables, we have shortened the responses for visual reasons. Table 1 shows the full details of what the respondents were asked.

Table 1: Mapping between questionnaire statements and report statements

Long statement used in the survey questionnaire	Shortened statement used in this report
We will considerably increase our use of big data in the next 12 months	We will considerably increase our use of big data in the next 12 months
We use an accelerated data warehouse or database appliance (e.g. EMC Greenplum, Oracle Exadata, Teradata Aster Data, etc.)	We use an accelerated data warehouse/appliance
Users of a production system at my organization have access to big data (whether via a feed into an existing production system, or a big data-specific production system)	Users of a production system at my organization have access to big data
We use big data to integrate social information	We use big data to integrate social information
We are testing or using a big data solution in the cloud	We are testing or using a big data solution in the cloud
We are investing in new storage technologies in order to keep up with data growth	We are investing in new storage technologies
We are currently building one or more pilots in Hadoop	We are currently building one or more pilots in Hadoop
We are using Hadoop in production	We are using Hadoop in production

- Figure 1 below shows the results from this survey question.

Figure 1: Big Data in Europe in 2013



Source: IDC European Software Survey 2013, n=700
Countries included: France, Germany, Italy, Spain, U.K., Netherlands, Nordics
Agreement with a series of statements about Big Data
Note: Sorted by Strongly agree + Agree

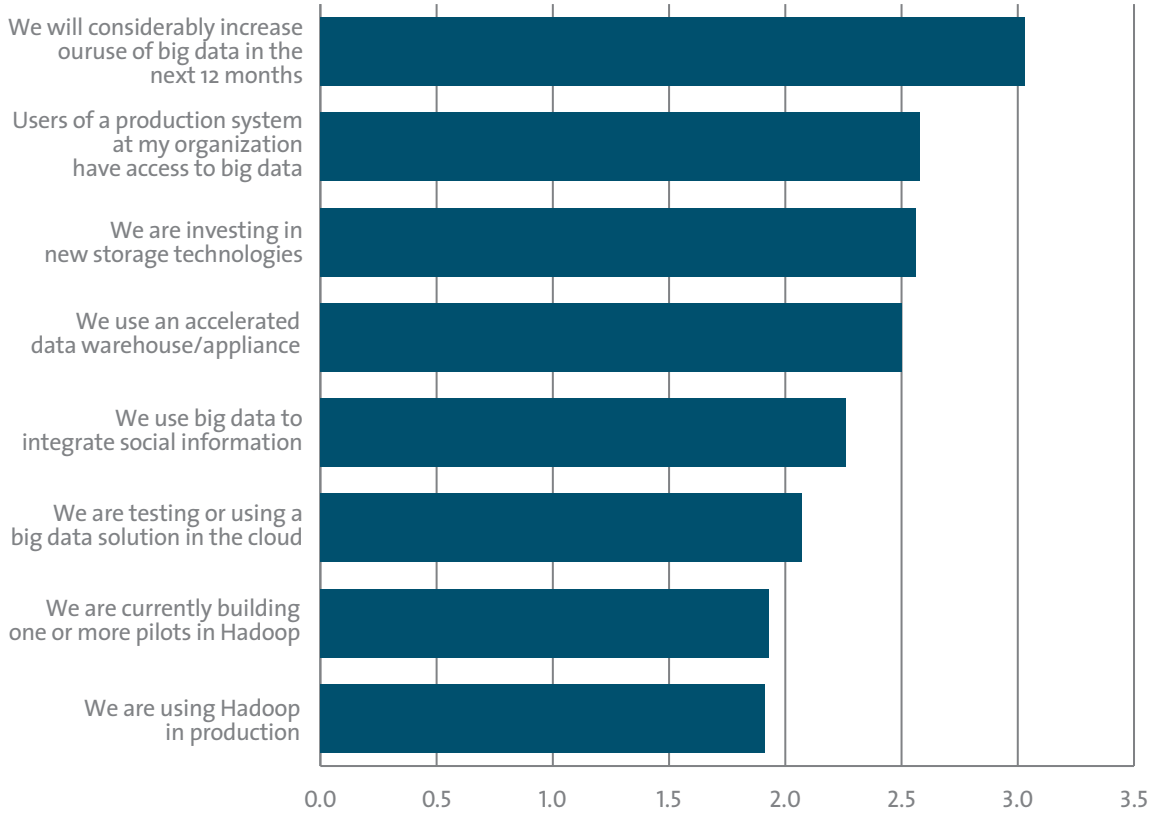
- Figure 2 shows the results when averaged across a 5-point Likert scale, so Strongly Agree = 5, Agree = 4, Neither Agree nor Disagree = 3, Disagree = 2 and Strongly Disagree = 1. Let's look at the responses to each statement in detail.

We will considerably increase our use of Big Data in the next 12 months – 40% agreement

- When looking at agreement with Big Data related statements, the most popular answer is an intention rather than an activity, which is to considerably increase the use of Big Data in the next 12 months. 40% of respondents answered "strongly agree" or "agree" to this question.

- In 2012, 25% of organizations stated they would considerably increase their use of Big Data in the next 12 months. However, in 2013, only 21% of organizations say their users have access to production Big Data, whether via a feed into an existing production system, or a Big Data-specific production system. The 2013 response of 40% indicates that a somewhat higher proportion of organizations intend to considerably increase their use of Big Data than last year – but to say all of these will follow through successfully on these intentions, and that these intentions will result in production deployments, would be overly optimistic.

Figure 2: Big Data in Europe in 2013 - Average Scores



Source: IDC European Software Survey 2013, n = 700
Countries included: France, Germany, Italy, Spain, U.K., Netherlands, Nordics
Average results across a 5-point Likert scale, so strongly agree = 5, agree = 4, neither agree nor disagree = 3, disagree = 2 and strongly Disagree = 1.

- We believe, taking a cautious approach that the 2013 survey results indicate that by the end of Q1 2014 around a third of European organizations will provide access to Big Data in production.

We are investing in storage technologies to keep up with data growth – 27% agreement

- This is a small increase on last year's survey responses, when 25% of respondents said they were investing in storage technologies to address data growth. Data growth is not a new phenomenon for IT managers however, while data growth in transactional systems is largely under control, organizations are struggling with an explosion of unstructured content. Such content can vary widely - from high definition digital images taken during product testing that need to be kept for product liability reasons, sensor-based information in energy grids, call records and positioning data from mobile phones, MRI scan images in health care and many more. The storage technologies that are typically invested in to cope with this explosion are scale out file systems, archiving products and, to a lesser extent, object storage.

- Storage is also starting to become a concern when customers are evaluating Hadoop distributions. For storage, Hadoop uses its internal Hadoop Distributed File System (HDFS) component, which is not considered an enterprise class file system. Hadoop distributions like MapR have improved the storage component to make it enterprise-ready. Another approach is to replace the Hadoop file system with an enterprise-class file system like Symantec's Veritas cluster file system, or RedHat's cluster file system.

- Performance is also a concern in Big Data environments, as storage infrastructures are challenged by the demand for high speed data ingestion and real-time analytics. Moving storage and compute closer together and performing the compute directly on the storage system are some of the approaches used to solve the performance problem. Currently, Big Data requires direct attached storage to guarantee the required performance of the storage system, in the medium term however, organizations will want to run Big Data workloads on a shared storage infrastructure with quality of service features enabled, in order to ensure optimal utilization of the storage infrastructure.

Users of a production system at my organization have access to Big Data, whether via a feed into an existing production system, or a Big Data-specific production system – 21% agreement

- This response was in third position when ranked by percentage agreeing with the statement (21% agreed) but in second position when ranked by average. This shows there is some polarization in the responses: a higher level of agreement but also a higher level of disagreement. This is partly because there was a higher level of disagreement with the statement „we are investing in storage technologies".

- This result emphasizes that Big Data is an evolution from business analytics. Remembering that only 8% of end user organizations are embracing Hadoop, the fact that 21% of users have access to Big Data indicates that European organizations also struggle with other Vs than variety. Big Data is not only about Hadoop and the integration of social data.

We use an accelerated data warehouse/appliance – 21%

- This relates to the previous question and again underlines the evolutionary nature of much of Big Data – organizations giving access to data from accelerated data warehouses count as giving their end users access to Big Data. Accelerated data warehouses are somewhat mature but they are also not traditional in the sense that an Online-Transaction-Processing (OLTP) database is, so they are classified as Big Data. Rates of accelerated data warehouse usage vary across industries, with strong take-up in B2C industries that need to handle consumer data, such as telcos, retail and finance.

We use Big Data to integrate social information – 17%

- This is a respectable, although not high, total and is consistent with last year's results when 14% of organizations stated they were integrating social information by using Big Data technologies. So this is rising but not very quickly.

- We believe these results apply to organizations using Hadoop platforms, other big data architectures, and organizations with subscriptions to social integration services. This emphasizes that Big Data technologies will often not be developed in-house but will instead underpin cloud services and other applications.

- The ability to use Hadoop to integrate information from social networks is a huge step forward in terms of gaining value from information. Traditional tools and techniques allow structured data to be integrated and analyzed, but social network information challenges those tools in many ways – lack of specified data format, various levels of completeness, etc. However, not all organizations want or need to set up a Hadoop system to integrate social data – many will use simple cloud services or applications that offer social dashboards, alerting, and time line views. Some such social data integration companies also offer consulting services around social strategies, while PR and media agencies who are already working on marketing and social strategies often offer social data visualization as part of their services.

We are testing or using a Big Data solution in the cloud – 11%

- Business analytics was somewhat late to the cloud, because when integrating transactional data from inside the organization, into sensitive aggregated data sets, the cloud did not particularly help with the challenges that this brought. However Big Data is drawing business analytics to the cloud, via a specific set of use cases:

  - Analyzing cloud-based data. As more organizational data has moved to the cloud, such as CRM system data, more business analytics has moved to the cloud. It is natural for analytics to run close to the largest or most important source of relevant data – we call this effect "data gravity".

  - Analytic sandboxing, where users run exploratory routines to investigate the content and value of the data. Until the value is proven, the system is temporary. This could be at the IaaS, PaaS or SaaS level, with the customer using analytic processing on demand, infrastructure on demand, or pay-as-you-go style system rental.

  - Some Big Data production processing requires significant peaks in infrastructure, for which the cloud is ideal. An example from the UK is a mobile application for voting and predicting results in a Saturday night TV talent show – the application runs all week allowing people to listen to the participants and make their predictions, but a vast peak of use occurs on Saturday nights during the TV broadcast.

  - Data that is too large to integrate as a whole or data that requires filtering or preprocessing before integration. The clearest example here is trying to integrate social data. Collecting all the social data that may be relevant and analyzing it in house will have a high impact on data volumes – it makes sense to run analytics or simple filtering to reduce data volumes before putting it into your own system.

8% of European organizations are working with Hadoop: 4% are piloting in Hadoop, and 4% are using Hadoop in production

- We believe this is an accurate figure. Hadoop is a relatively nascent platform and adoption in the US has significantly outpaced that in Europe. However, Europe is starting to catch up. Enterprise Hadoop vendors MapR, Cloudera and Hortonworks have all entered the European market in the past 2 years.

1.5 Evolution between 2012 and 2013

- Figure 3 below shows the evolution of the statements we asked in 2012 as well as 2013.
- The level of agreement with "we will considerably increase our use of Big Data in the next 12 months" was 25% in 2012, and has increased to 40% in 2013, demonstrating a wide increase in intention to undertake projects.
- Agreement with "we are investing in new storage technologies to handle data volumes" increased modestly from 25% to 27%. It may be that organizations are focusing more on the analytical layer of the Big Data stack, as that is what relates directly to business value.
- Agreement with "we are using Big Data to integrate social information" has also modestly increased from 14% to 17%. This shows that organizations that were strongly focused on integrating social data did so in the earlier days of Big Data.

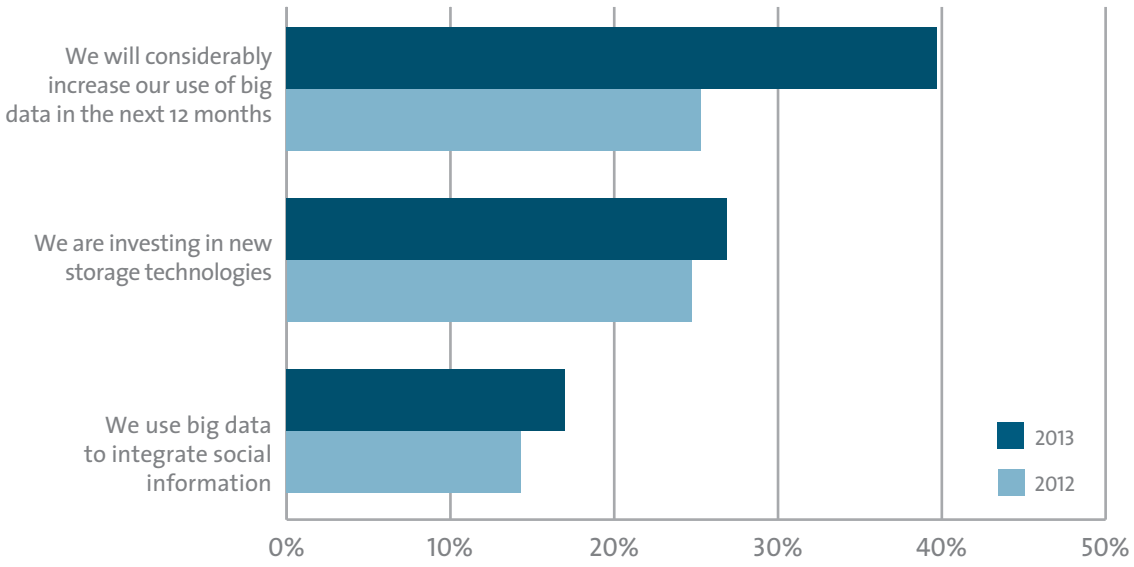
1.6 Industry Differences around Big Data

- Big data is driven by specific use cases rather than applying the technology across the board. Certain industries, particularly business-to-consumer (B2C) industries, are naturally geared towards these. This is because B2C industries have long been aware of the challenge and opportunity relating to being able to deduce insights from their transactional datasets. This section shows data for the most significant industry sectors.

Big Data in Financial Services

- Financial Services is an advanced sector in the adoption of Big Data in Europe, and IDC believes it will remain so. The sector has by far the highest proportion of organizations intending to increase their use of Big Data at 92%.
- The sector also has a higher rate of deployment of Big Data to end users (35% of Financial Services organizations compared to 21% of all respondents). This builds on the maturity of the sector in traditional business analytics.Large data sets and the need to process streaming data have pushed many applications into the region of Big Data in terms of volume and velocity.

Figure 3: The Evolution of Big Data in Europe - 2012 to 2013



Source: IDC European Software Survey 2012 and 2013
Note: Strongly agree/Agree responses in 2013 (n=700) and 2012 (n=700)

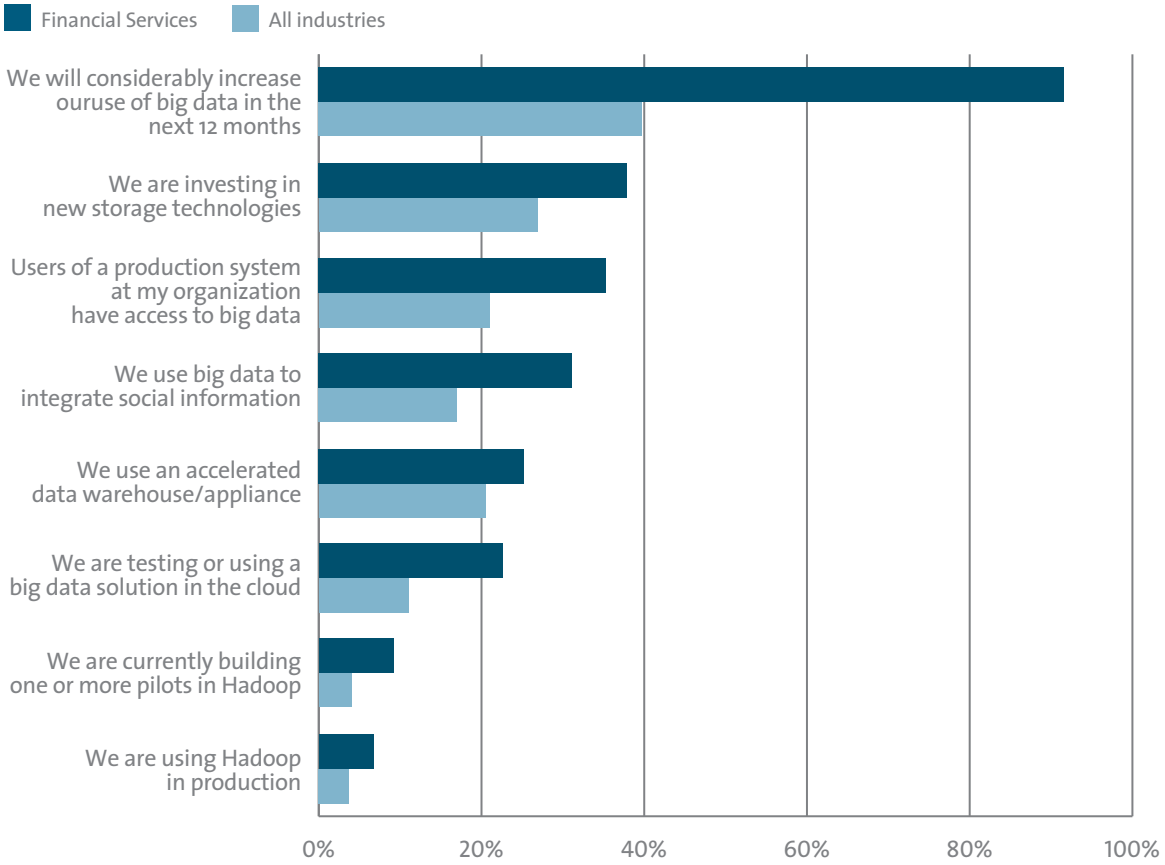
- Although variety has previously been of less interest, in this sector 9% are using Hadoop in production and 7% are piloting in Hadoop, compared to 4% and 4% respectively across all respondents. The sector has been quick to adopt Hadoop, due to the "build rather than buy" mentality that is more prevalent in Financial Services that in some other industries – these organizations often have access to developer teams and a natural inclination to investigate new technologies to see what their ramifications could be.

- 38% are investing in storage (27% of all respondents) showing that a high proportion of Financial Services organizations are focused on handling large, and growing, volumes of information.

- Financial Services has a noticeably high level of integration of social information, and this has dramatically risen since 2012, when only 15% of organizations in the sector agreed with this statement. Social data is important to banks as a warning mechanism to alert them to occurrences such as banking outages, security issues, and customer concerns. When the UK operation of a global bank suffered its second service outage in 6 months, the bank's corporate communications department first found out about the outage via reports from customers on Twitter – before the IT department had found the reason for the outage. This allowed them to respond more quickly to customers by formulating a statement without needing to wait for their IT department to officially confirm that there was a service outage.

- Figure 4 shows the survey responses for the Financial Services sector.

Figure 4: Big Data in Financial Services in Europe in 2013



Source: IDC European Software Survey 2013; Note: Responses of Strongly agree or Agree; All Industries n=700, Financial Services n=119

Big Data in Telco & Media

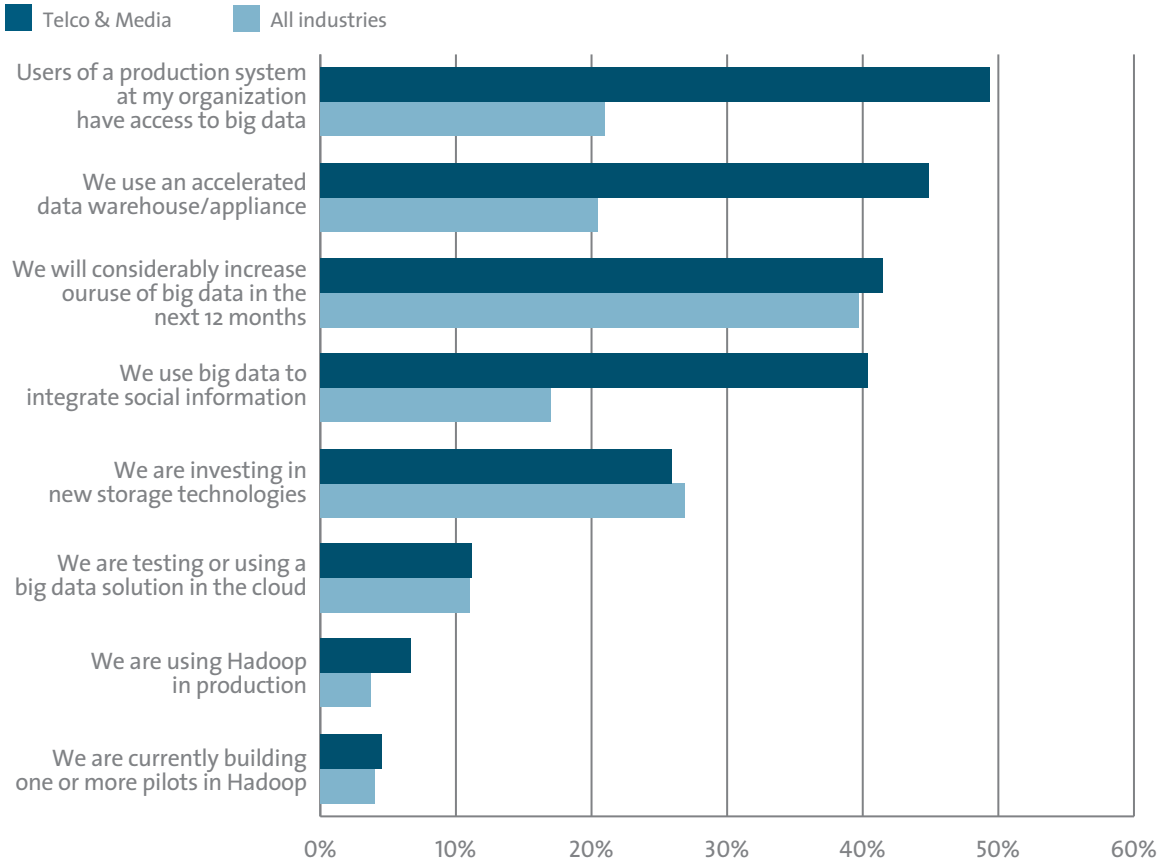
- The telco & media sector is unique in that the highest proportion of responses came for activities rather than intentions. All the other sectors had the highest proportion of responses to "We will considerably increase our use of Big Data in the next 12 months." For the telco & media sector however, there is an existing level of maturity and activity, mainly focused from the telco side on ingesting and analyzing call detail records (CDRs) in order to support applications such as churn reduction, revenue assurance, marketing analytics, network analytics, and customer experience management.

- However, telco & media is the third industry when ranked in order of expectation to increase use of Big Data in the next 12 months, coming in close to average with 42% agreement (the average was 40%). IDC believes this could indicate that organizations in this sector are not hungry enough! Because of their focus on gaining value from structured data, they may miss the broader variety-based and velocity-based Big Data opportunities.

- So what is the potential for Big Data in telcos? The maturity could indicate analytic sophistication, and generally the more analytics an organization uses, the greater it's analytic focus, therefore there is an ongoing driver to use more analytics.

- Figure 5 below shows the survey responses for the Telco & Media sector.

Figure 5: Big Data in Telco & Media in Europe in 2013



Source: IDC European Software Survey 2013; Note: Responses of Strongly agree or Agree
All Industries n=700, Telco & Media n=89

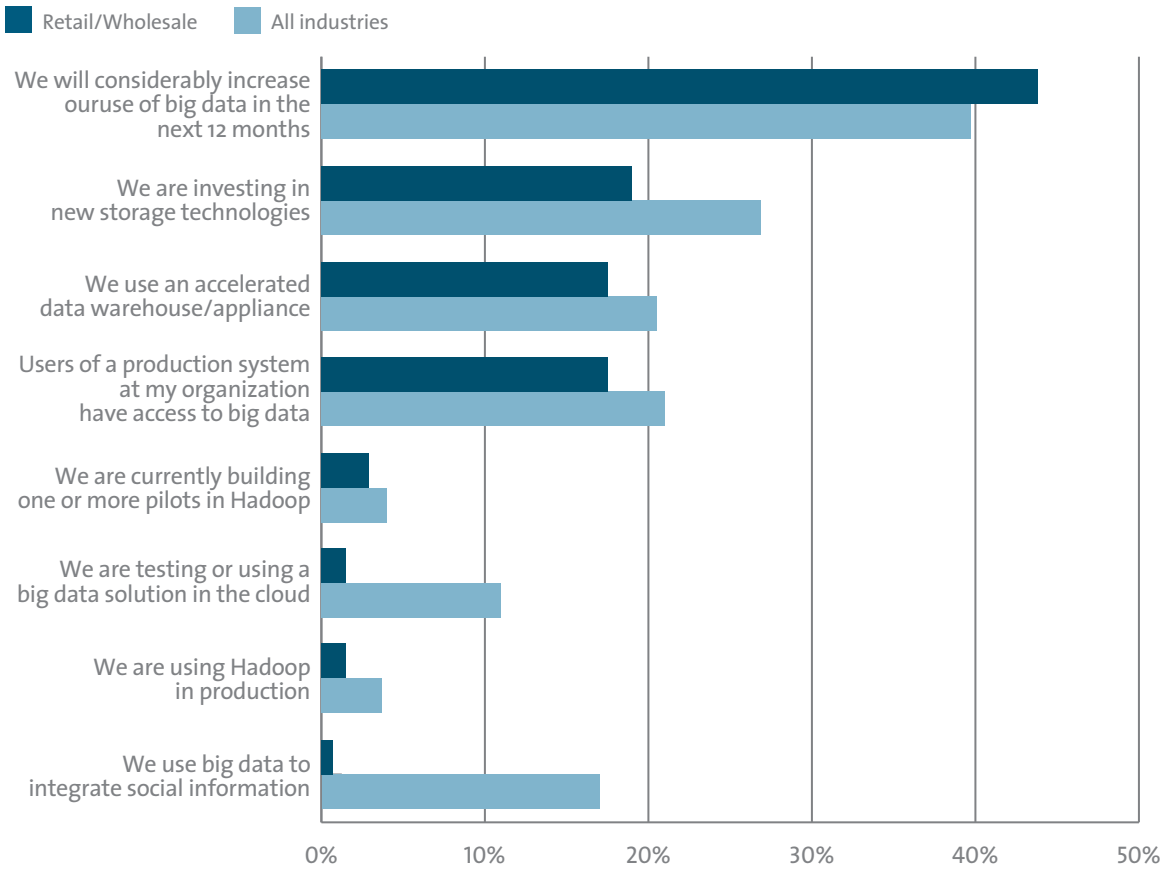
Big Data in Retail

- Retail had a surprisingly lackluster response to the survey, with the industry's maturity in analytics not really translating to Big Data. Although the industry slightly outpaces the average in responses to "we will considerably increase our use of Big Data in the next 12 months" all the other responses are below average. One that is particularly surprising is the low rate of retail organizations saying they use Big Data to integrate social information, when retailers are at the forefront of gaining business value from social data. From this we conclude that although retailers are definitely integrating this data, they would not describe what they are doing as relating to Big Data.

This is probably because rather than building their own Hadoop clusters and integrating data themselves, they are working with external providers of social media dashboard and marketing analytics functionality.

- Figure 6 below shows the survey responses for the Retail sector.

Figure 6: Big Data in Retail in Europe in 2013



Source: IDC European Software Survey 2013
Note: Responses of Strongly agree or Agree, n=700

1.7 Big Data in the Enterprise v SME

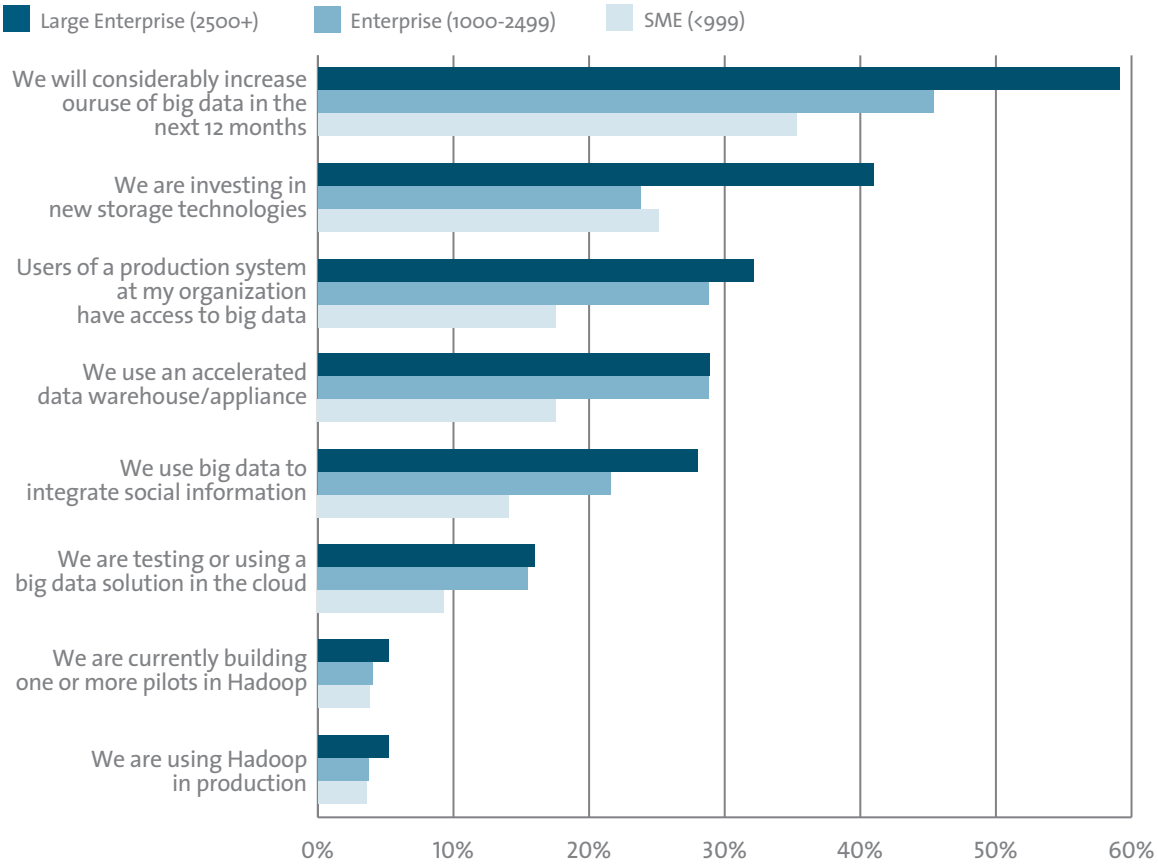
- Figure 7 below shows the agreement with the Big Data statements split between SME (between 50 and 999 employees), enterprises (1000-2499 employees) and large enterprises (2500 employees and over).

- Generally the large enterprises lead, particularly with the expected increase in Big Data activities during 2013, investment in storage technologies, access to Big Data and integration of social information. This segment also leads in Hadoop but by a surprisingly narrow margin. We would expect the largest enterprises to have the highest levels of implementation and activities around Big Data because they have the largest problems with data volumes, fragmented data, fragmented system landscapes and a broad population of end users, all of which mean they adopt business analytics tools and techniques earlier than small organizations.

- The enterprise segment falls between the large enterprises and the SME in terms of intended activity in 2013, and end user access to Big Data in production. The level of adoption of accelerated data warehouses/ appliances is the same as for the large enterprises.

- The SME segment has surprisingly high rates of Hadoop adoption: this could be due to new businesses, particularly in the technology sector, that are driven by the availability of Big Data technologies to drive innovation. The video game development sector is a great example of this.

Figure 7: Big Data in Europe in 2013 - Large Enterprise v Enterprise v SME



Source: IDC European Software Survey 2013, n=700
Note: Responses of Strongly agree or Agree

2 Future Outlook

IDC believes market growth in Europe will proceed at a slower rate than from North America over the past few years but the region will start to catch up.

- **IDC expects a strong ramp up in adoption of Big Data in Europe over the next year.** 40% of respondents expect to considerably increase their use of Big Data over the next 12 months. When compared with the same data for last year and current adoption rates, we believe these responses indicate that by the end of Q1 2014 around one-third of European organizations will provide access to Big Data in production.

- **The number of Hadoop systems is currently small but penetration will grow.** 8% of European organizations are working with Hadoop: 4% are piloting in Hadoop, and 4% are using Hadoop in production. Hadoop's ability to integrate vast amounts of variable format data is immense: but as a nascent open source framework requiring rare developer skills, Hadoop is not for everyone. Vendors including Cloudera, Hortonworks and MapR are entering the European market.

- **Maturity drives adoption.** Greater maturity in analytics means more understanding of the benefits of Big Data so should drive adoption. However, organizations with understanding of the challenges of integrating internal transactional data are likely to feel quite fearful of the highly complex world of Big Data, which moves away from the "single version of the truth" residing in an enterprise data warehouse. Some organizations will be overtaken to success in Big Data by their competitors who did not particularly embrace business analytics as they realized it would never help them gain value from specific text-based sources.

- **Big Data is going to be a big boom for the IT industry.** Web sites that gather significant data need to find ways to monetize this asset. Data scientists must be absolutely sure that the intersection of disparate data sets yields repeatable results if new businesses are going to emerge and thrive. Further, companies that deliver the most creative and meaningful ways to display the results of Big Data analytics will be coveted and sought after.

- **Trust will be a significant factor in shaping the Big Data market in the medium- to long-term.** The full impact of the 2013 revelations about NSA mass-surveillance of data via interception, and also via agreements with some large US technology vendors, is not yet fully clear. However, it is likely that European organizations will become more determined to store their data where it is subject to European data protection legislation, and the pressure on US organizations to meet this need will increase. Trust between IT suppliers and their customers is becoming more important as Big Data is more widely adopted, because great insights bring great power.

3 Appendix

Methodology and definitions

The methodology used to collect the data and data definitions are published on our website www.eito.com/method-reports-download.

Terms and conditions

All terms and conditions published on the EITO website apply: www.eito.com/termsandconditions.

Customer Feedback

We are always interested in customer feedback so please let us know which parts of the report you find particularly useful (data tables, analysis of key trends, country analysis, etc.) and which are of less interest to you. Send us an email at service@eito.com.