Big Data Challenges and Trends

Hafeez Niazi
School of Computer and Information Science,
University of South Australia, Mawson Lakes, Australia

Abstract: Big data is becoming the hottest topic and flash discussion point in any format of information media today. Most of the largest organisations linked to information technology investing in and advertising about it, like Oracle, IBM, Google, Yahoo and SAS. Big data is about data which can be in whichever format and size, utilize by the organisations without losing a single piece of information. Big data is not simply about the size but also about the content and data which will make organisations business more responsive and dynamic. Every day data is growing and expanding in different formats to accomplish and proficiently process rising thousands of petabytes data each day, current technologies and method cannot fulfil the necessities and needs new trends of technologies like 'big data'.

Key words: Clouds, big data, business intelligence, data warehouse, data over the internet

INTRODUCTION

Data processing for diverse types of data on big cluster delivers more proficient performance as associated with various database management systems. The enormous amount of data comes from different sources in various kinds of patterns such as videos, pictures, social media sites and sensors, RFID, purchased and transactions, cellular phone and satellites signals. Russom discussed that if business needs to know what's happening in their business, they need complete and thorough level of information. The source of information coming from sensors, devices, web applications and social media which is never been exploit for Business Intelligence (BI) purposes before. Hopkins (2011) described that inundations of information is coming from internet, sensors and images. The information which is thousands times higher data rates that IT systems managing currently which is causing the business losing valuable information. The companies are demanding for more improved and detailed value of information. Information is the only assets that companies earn and run their business on, if that information is out dated or lost it will bring the whole organisation down. Nasar and Bomers (2012) discussed the idea of big data as huge dataset which comes from conventional and non-conventional sources. Financial industry is delicate to data concerns because it always needs reliable and precise information about financial data and customer information. As discussed by Valle *et al.* (2011) that Big data is becoming the need of every industry in the world, new technologies are retrieving and collecting more data

than ever before. Each industry wants to utilize every piece if its valuable information to strive and excel with its competitor better than others.

Big data is three vs: Russom described that big data is more than about volumes which involves variety and velocity of organisation's data. Oracle and IBM also share the same concepts about three Vs but oracle goes one point ahead with four Vs with the addition of value in its v definitions. When it comes to data, every organisation is different from others and they have different requirements to store and present data to business.

Volume: This sort of non-traditional data is generated by machine in enormous quantities.

Velocity: This sort of non-traditional data is generated by social media.

Variety: This type of data is traditional data and modified gradually.

Value (extra vs added by oracle): There is always valuable information is concealed among huge form of non-traditional data; the task is to find that data for analysis.

According to Fig. 1, this shows volume, velocity and variety in details. Volume will contain physical existence of the records in the form of tables, records and transactions. Velocity will involve, real time and streams sort of data. Variety will have all sort of data including structured and unstructured.
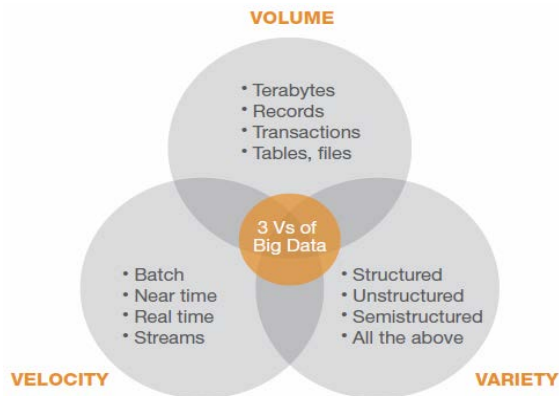
Fig. 1: The three vs of big data

Big data produce business significance: The Global institute (Gregory and Hunter, 2011) described that data volume is increasing 40% per year and it will rise forty four times at the end of 2020. This will not only the matter of managing volumes of data but the value of the data will also matters. The data is the value and main assets of the organisations.

Producing transparency: Big Data will make easy access to the applicable data for different stake holders in well-timed which can produce incredible value for the organisation. By making data availability at time and easy accessibility will drop search and processing time which will create value and customer satisfaction for the organisations. Fox (2012) discussed about the big data impact on the health care system where the false claim can be easily identified by individuals.

Presentation of inconsistency and increase performance: Big data will make organisation's transactional data available in in digital form which will allow them to collect precise and detailed information of data on real time. Nasar and Bomers (2012) discussed about financial intuition grieves from inefficiencies around data. It will save time and human efforts of manual processing financial applications. It will also provide the whole picture of customer relationship with other institutions to regulators authorities to analyse risk associated with it. Watts (2012) described that big data will deliver the great openings to the financial institutions to improve the customer satisfaction with the support to established the vigorous customer communication management program.

Segmenting to modify actions: Big data will permits organisations to generate segmentation of their data which will permit them to shape their products and services more accurately. Big data practices will allow real time micro-segmentation of customers to aim advertising and promotions.

Better decision making and computerized algorithms: Big data allows organisations to improve its decision making and display data that is hidden by computerized algorithms. It will help tax collecting departments and other financial institutions to banner candidates for further detailed analysis. With the help of big data organisations can analyse the entire data sets which will provide them better decision making against the real time data. Crump (2012) describe that big data will provide insurers organisation to analyse whole data sets to enhance business process around customers and products to improve customer service.

Inventing novel business models: Through big data companies will be able to produce different products and services, improve the existing one as well and develop innovative business models. It will allow the manufacturers companies to use the actual data to develop next generation's products and to produce after sales inventive contributions. Hamm described that companies wants to understand what sort of products customer really want and at the end of the day what they will buy all linked to bid data.

Big data and role of Enterprise Architecutre (EIA): Gartner described that old-style methods to EIA are obstructed by big data. EIA move its emphasis from optimization and standardization to different practices like coordination and externalization in enterprise life cycle. Big data shift the focus from data warehousing to further towards data pooling.

Deploying big data to enterprise: Rai and Oracle discussed about to bring big data into enterprise architecture, so the problem is find the relevant information and use it in real time. Organisations are looking for to make most of big data and use it within the existing infrastructure. Oracle come up with the big data model where it can deployed efficiently in organisation's existing infrastructure.

Acquire: In this phase, big data states to data streams of higher velocity and higher variety, this information should be capture with no latency with simple quires to manage huge volumes of data. NoSQL databases are used quite often to obtain and to accumulate big data. The data kept

Fig. 2: Divided solution spectrum big data; Oracle Big Data for the Enterprise in 2012
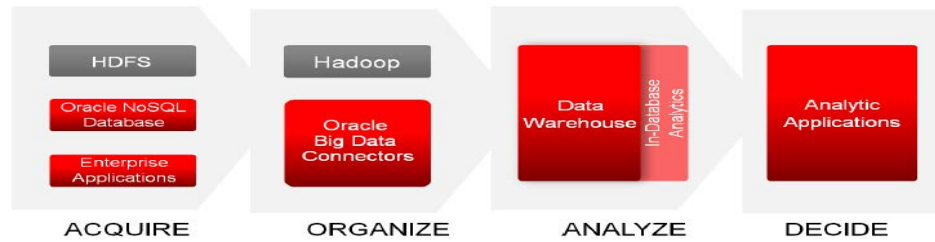

Fig. 3: Oracle's big data solution

Table 1: Key drivers to consider

Buisness drivers	IT drivers
Better insight	Reduce storage cost
Faster turn-around	Reduce data movement
Accuracy and timeliness	Faster time-to-market
	Ease of management and operation
	Security and governance

in NoSQL database is normally of a great diversity because the systems are planned to purely capture all data without classifying and analysing the data. Ghosh (2010) discussed about to use NoSQL data storage engines to exploit the storing data the same way it is stored. Some big companies still preferred to use the relational database and to measure data for all sort of kinds. Google considered BigTable and Amazon come with Dynamo to utilize all data form huge relational database.

Organize: In this phase data is stored on its original location, so it will save both time and money by not transferring huge volumes of data. This phase is more like integration of data in data warehousing terms. Apache Hadoop is the modern technology which will permits large data volume to be organised and managed while keeping the data into its original data storage cluster.

Analyse: In this phase, all analysis taken place on the data. The analysis of big data is essential to support crucial detailed statistical analysis and data mining on massive and different varieties of data types. It will

provide organisation to not only analyse new data but also keep the old analysis to deliver an innovative viewpoint on old problems.

According to Fig. 2 which shows big data spectrum with NoSQL and SQL perspective, NoSQL is required on larger data sets to acquire all of the data in timely manner on the other hand SQL is used under more secure and administered environment against relational database systems.

According to Fig. 3 which shows the complete oracle big data solutions, where Hadoop Distribution File System (HDFS) is used to collect and organise data and theses accumulated outcomes later loaded into relational DBMS system.

Key drivers to consider: Oracle discussed about the key drivers which will need to be consider when making a decision about big data. Key business drivers list shown in Table 1 given.

MATERIALS AND METHODS

Big data challenge or oppertunity: Arnold (2012) elaborate more about the challenges involved in the implementation of big data as organisation go further along with it. One of the main challenges involved the lack of staff with big data and analytical skills and second one it's not limited to local contexts, it's more global. Matelski (2012) discussed that big data contains several prospects but dealing with fast progresses enhance challenges, budget and complications.

Russom discussed organisation's thinking about big data as problem or an opportunity, only 30% reflect that big data as problem because only large volume of data is not only the solution for some organisations. On the other hand, 70% of the organisation sees it as a prospect to thorough studies of their data which contains new evidences about their consumers, associates, costs, processes and procedures which can be used for business proficiencies.

Big data getting developed and established: EMC explains that many companies are utilizing different sources to accomplish superior consideration of their, customers, associates, employees and processes. Companies are exploring and processing more than organized and transaction based data which contains videos, social networking, RFID logs, sensor networks, search indexes, environmental conditions and medical scans.

Universal methodology: Hopkins (2011) discussed to develop strategies for big data to capture the right metadata which will be analysed, maintained and store quickly. It will provide the organisations to understand the large data to design business intelligence solutions which can later link to build the strong Service Oriented Architecture (SOA).

Variety of digitization sources: Valle *et al.* (2011) discussed about variety of digital sources which big data can support and organisation can fully utilized different pieces of the detailed information from different sources. Big data supports different type of sources like media and entertainment, healthcare, life sciences, video surveillance, transportation and logistics.

It will help organisation's those trading globally which will provide the organisation's to make their decisions, efficiently and well-timed based on their transactions.

Big data governance issues: Adrian and Kathy described that customer data is the precious assets for the organisations which required vigilant governance and protection. Managing big data is not easy and it will require lot of precaution and architectural framework to make sure the quality and security of the underlying data. Tankard (2012) discussed about the data governance issues related to more towards security. Big data provide the detailed level of analysis and sensitive information to trade secret, financial archives and knowledgeable belongings. Make information available centrally will provide informal and treasured target for attackers or

hackers which can harm the status of the organisations. Matelski (2012) also discussed about big data issues like data governance problem in public and private organisations. To address these sort of issues organisations needs to re-design its data policy which supports and simplify data inventions. As explained by Agrawal *et al.* (2011a, b) that data governance is great worry for records which is coming from multiple sources. Data gathered from different geographical locations where user needs to share the locality with the service provider which will cause the countless privacy issue. There is still great concern in terms of big data, how to restrict access to information and what types of data which will end up with a lesser amount of detailed information. In practical world data is not still and it is always changes, its volume, velocity and variety, so fundamental procedures will not gone work in this scenario.

EMC (2012) discussed that big data technologies are still developing which will lead the companies to revisit, its privacy and security policies. Organisations are struggling to combine data in new ways which is not only causing the great security concerns but also how the data should be processed and where to keep that data? If companies will decide to use third parties tools to integrated the data whether it will be feasible and safe to do that or not? Data governance is still the main problem, who owns the data and who will held accountable if information is breached or how to track where information is stolen and by whom? Smith *et al.*, (2012) discussed about the massive obstacle in the implementation of the bid data is accountability of the data. Organisations should make resolutions about who will own big data and data sets. Organisations need successful governance strategy to implement data governance; otherwise it will not be durable to link unrelated data or information back to the client, customers and production tables to make right decisions at right time. Rajan the enormous challenge face big data in term of governance is the lack of budget, resources and awareness from the top management. Companies need proper data governance policy to make successful execution of big data, lack of procedures and direction principals around data causes governance issues for big data. In short data governance issues revolve around, data privacy, data ownership, governance policies, budget and interest from the executive level. Data governance is not a project or program, it is an ongoing process which cannot be stop at any time, even after the successful implementation of big data.

Big data and current trends: Big data is becoming the hottest topic and flash discussion point in any format

Fig. 4: Dimensions of Big Data

of information media today. Most of the largest organisations linked to information technology investing in and advertising about it, like Oracle, IBM, Google, Amazon, Microsoft, Yahoo and SAS. Big data is about data which can be in whichever format and size, utilize by the organisations without losing a single piece of information. Big Data is not simply about the size but also about the content and data which will make organisations business more responsive and dynamic. Every day data is growing and expanding in different formats, to accomplish and proficiently process rising thousands of petabytes data each day, current technologies and method cannot fulfil the necessities and needs new trends of technologies like 'big data'.

Mark and colleagues discussed about data volumes are increasing daily which only involves transactions data and other conventional data types. The next generation operation technology is part of that data which involves

audio, videos, streaming data and social network. Storing massive amount of data is the issue but doing the analysis on it also complex.

According to Fig. 4 shows different dimensions of big data, in quantification it will look for volume of data, variety of data formats and velocity of data. In access enablement and control, it will look for classifications, contracts and technology enablement. In qualification and assurance, it will look for fidelity, linked data and validation of the data. Russom described about the servery which held under TDWI shows different organisation's commitment and prospective growth.

According to Fig. 5, different level of commitment and potential growth shown by different participants, the way data is handled and process. Diverse big data trends is assembled according to Fig. 5 and also discussed by Walsh and Rodier.
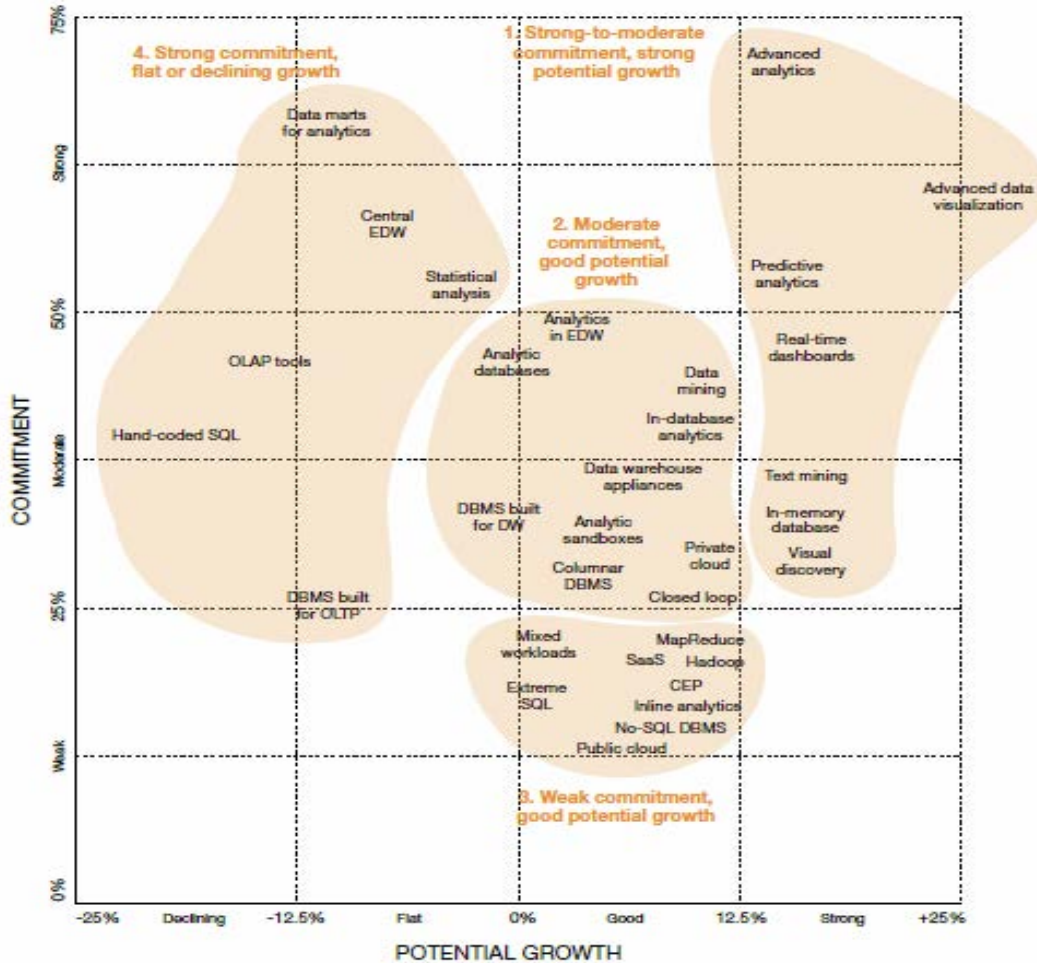
Fig. 5: Commitment vs potential growth

RESULTS AND DISCUSSION

Advanced analytics: It is the combination of different procedures and tools categories which includes prognostic analytics, data mining, statistical analysis, complex SQL, data visualization, artificial intelligence, natural language processing and database approaches which maintain analytics.

Visualisations: Big data provides Advance Data Visualisations (ADV) which characterizes millions of data points. ADV supports different sets of data varieties and data structure which is not adjustable into computer screen.

Real time: Organisation can observer and measure different business processes more regularly because of the bid data tools and analysis.

Structure and unstructured data: Organisations are handling unstructured and structured data more efficiently than ever before. This provides organisations improved data mining and fraud finding techniques to process under layers of organisational data.

Hadoop Distributed File System (HDFS): Big data is varied in terms of data types, HDFS will help to convert these different data types into data structure so it can fit into the traditional data base management system.

MapReduce: It will provide distributed parallel processing for different types of data collections which will make analysis easy on big data.

Complex Even Processing (CEP): CEP is used to grip big streaming data which is somewhat real time driven. Organisations are using event processing as the part of their data integrations.
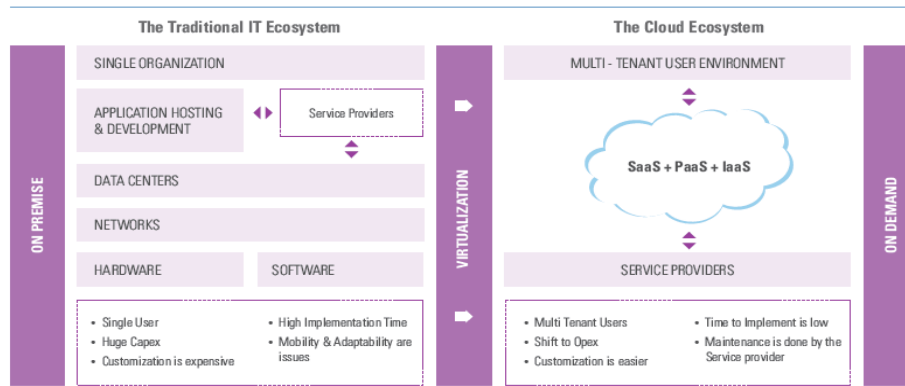
Fig. 6: Traditional IT VS cloud (Janet and Bradley, 2012)

No-SQL: No-SQL is used to handle information which is not relational, it will convert the information in tabular structure so it can be read-able by the relational database.

Data association, restructuring and purging: Big data will provide organisation's accurate analysis for the factual data which can track it back where it reside in data structures of the database.

Hardware for big data: Organisation needs to buy more hardware to analyse extraordinary datasets of structured and unstructured data which involves processing, storage, clustered servers and input and output speeds.

Security: Big data analysis brings security concerns as well because developers, analytics and numerous programmed service will be accessing data which will need proper access control, verification, data encryption, intrusion deterrence and auditing.

Antique information: Companies will be able to analyses historical data containing different time periods which will provide the lowest level of information for analytical models, trading and forecasting.

Governance and risk management: Transparent and lowest level of information will provide organisation better governance, auditing and risk reporting. It will also help the organisations to expand its present master data management policies.

Segmenting to modify actions: Big data will permit organisations to generate segmentation of their data which will permit them to shape their products and services more accurately. Big data practices will allow real time micro-segmentation of customers to aim advertising, consumer behaviour and promotions.

Cloud based infrastructure: Cloud based infrastructure will be needed to support local and global analysis, economic and business advancement openings in India, China and Brazil.

Reveal the worth of data: Big data analysis will provide organisations to un-reveal the hidden data in their processes to shrink cost and find out innovative prospects.

Extract, transform, load: ETL processes needs to redesign for big data structures to handle rising capacity of information.

Cloud computing and current trends: Cloud computing provides organisations the flexibility to use third parties services to access its information. There will be no infrastructure and management cost because organisation will be using service provider data centre and storage services to access to its own applications. Couture and Zaffos explained about cloud computing, company will save money when data is kept on the cloud relatively on the sites. The cloud appliances provide the facility to assimilate on sites and cloud storages.

According to Fig. 6 which shows cloud computing attributes compared to traditional computing. The collective cloud computing service contributions consist of elementary ranges.

Infrastructure as a Service (IaaS): This service provide infrastructure as a service, in which company outsource the hardware, networking gears, servers and storage.
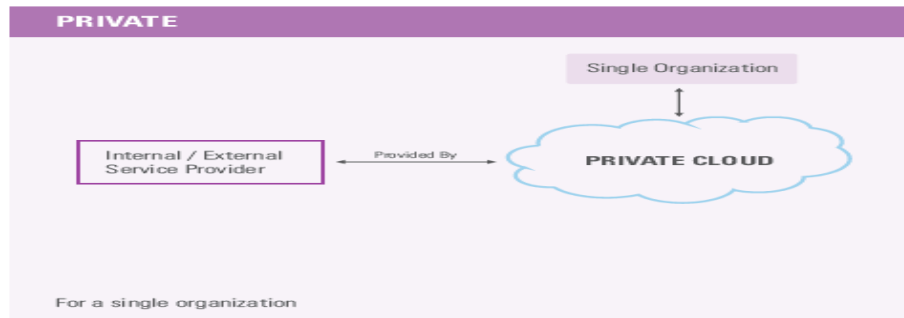
Fig. 7: Private cloud (Janet and Bradley, 2011)

Companies like Microsoft, Amazon and Telstra and other service providers delivers its personal equipment and responsible for administration, monitoring, marinating and running it.

Application platform as a Service (ApaaS): This service provides the whole application as a service, in which company lease the hardware, networking gears, servers, internet and storage. Companies like Google App Engine, Force.com and appspace to charge virtualise severs and related facilities and services for developing and running applications.

Software as a Service (SaaS): This service provides the software application as a service, in which company uses the application over the internet. Companies like Google, salesforce.com and other application service providers deliver OneSource and licenced soft wares and applications.

Service Oriented Architecture (SOA): This service provides the single function or partial software application as a service, in which allows companies to maintain and support communication between two different computing services or entities. Communication between two commuting entities and services are independent of each other, like Amazon Flexible Payment and PayPal.

Delivery models of cloud computing: Janet and Bradley (2011) discussed about different delivery models of cloud computing which includes Public, Private and Hybrid Cloud.

Private cloud: Private cloud deployed on company's own site or datacentre belongs to the service provider. The private cloud makes logic to be deployed and contain within the zones of company infrastructure which will provide more security and control over the company's

infrastructure. It will provide access control to companies which will assured certified access to infrastructure and its applications, hardware utilisation and assets management. Vendor hosted or partner allowed the companies to move their applications to the service provider secure data centres. Companies will be saving cost rather than investing and buying the new technology not only equipment but also release pressure on in house resources (Fig. 7).

Public cloud: Public cloud allows organisation to utilized infrastructure and applications through the internet. Organisation's end users will be accessing applications, servers, networks and services on demand through web browser on their laptops or workstations. Companies no longer required to buy the expensive equipment, when applications and services are accessible through internet on little cost. The main problem moving with public cloud is security and regulatory issues (Fig. 8).

Hybrid cloud: Hybrid will provide the combination of both private and public cloud to the organisations. It will provide the organisations the choice to switch between private and public on peak processing times, when resources required over the network. The companies will deployed the application which are more important and essential on the private cloud behind the firewall, like less security services and applications can place on public clouds. Hybrid model will allows the companies to maintain and control peak time periods, save costs and security on precarious applications (Fig. 9).

Smith *et al.* (2012) discussed about cloud computing trend and how cloud computing will become the essential part of the organisation's infrastructure implementation in the future.

Hybrid cloud computing vital for organisations: Hybrid cloud will become essentials for the organisations in future because it will offer a single cloud solution with the
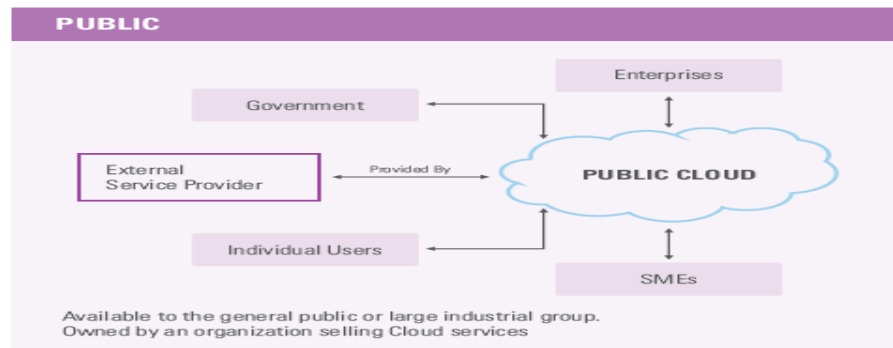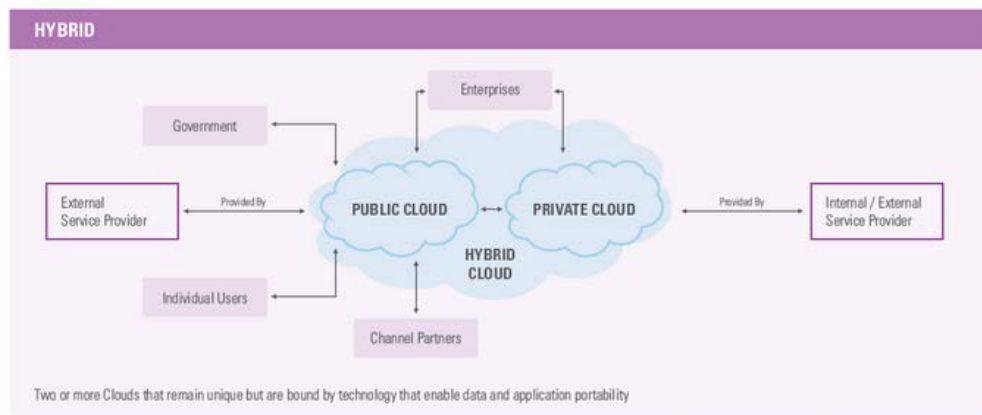
Fig. 8: Public cloud

Fig. 9: Hybrid cloud

mixture of public and private cloud computing services. Organisations will be concentrating more on data and application integration, relating exterior and interior applications with hybrid computing implementations.

Cloud computing service provider assistance on cloud consumption: Organisations will need the help of cloud computing service to liaise with IT department to discuss the purchasing and cloud solution options.

Cloud-centric deployment: Organisations will be focusing and exploring different options and prospects to transfer current enterprise workload on application infrastructure or cloud system. To make most out of the cloud model organisations need to plan applications deployment where most of the workload is vastly adjustable and scalable.

Upcoming operational models and data centre based on cloud model: In coming years most of the organisation will be designing and developing its own cloud computing mode to support its operational model and data centres.

Future developments, issue identification and analysis in company A: Company A is the enormous big mining giant which has several mines located all over the world. Each site has its own business rules and doing things according to their requirements which is not ideal if company want to implement a global and standardise solution. Company A is using Oracle stacks which is using Oracle Warehouse Builder (OWB), Hyperion Performance Suite, Business Intelligence Tool (OBIEE) and Ellipse as ERP system (Also based on Oracle) and oracle databases. The company is expanding quickly in each area which involves different subject areas like assets management, mobile assets, procurement, sales, commercial, planning, finance, HR, budgeting and forecasting. The result of that data sets are getting bigger and larger, transactions and operations are expanding enormously every day which need proper analysis, data quality and to make business operations uniform and standardise each site and location. A Company's corporate affairs and legal business functions and department are getting busier than ever before which needs to look for every media to get the right and proper

Table 2: Making Sense of Big data (Kimball, 2012)

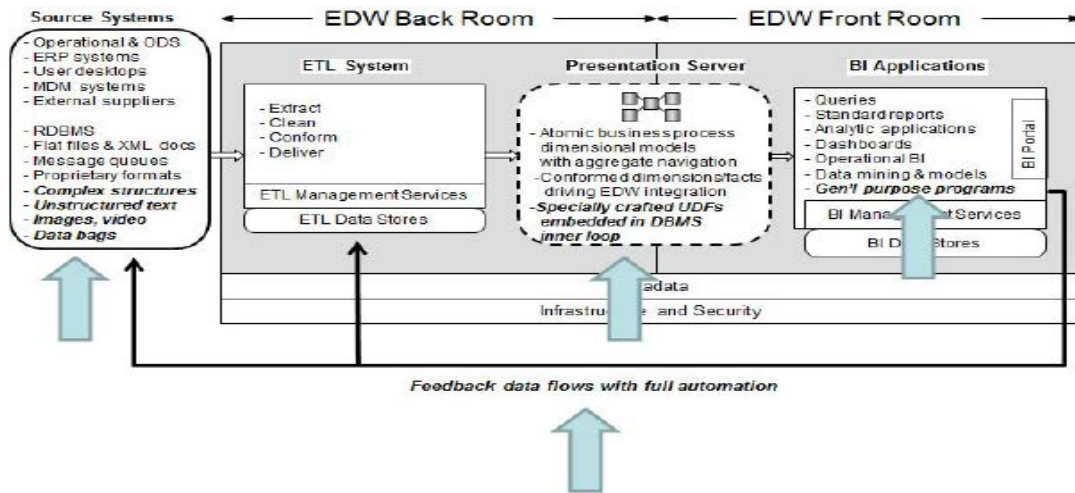| Factors	Vector matrix or complex structure	Free text	Image of binary data	Data "bags"	Iretative logic or complex branching	Advanced analytic routines	Rapidly repeated measurements	Extreme low latency	Access to all data requires
Search ranking		×	×	×	×	×			×
Ad tracking	×	×	×	×	×	×	×	×	×
Location and proximity	×		×	×			×	×	
Casual discovery	×	×	×	×	×	×			×
Social CRM	×	×	×	×	×	×		×	×
Document similarity	×	×	×	×	×	×			×
Genomic analysis	×	×	×		×	×			
Cohort groups	×	×		×	×	×			×
In-flight engine status	×		×		×	×	×	×	
Smart utility meters	×		×		×	×	×		×
Building sensors	×		×	×	×	×	×	×	×
Setelite images	×		×		×	×			
CAT scans	×		×	×	×	×			×
Financial fraud	×	×	×	×	×	×	×	×	×
Hacking detection	×	×	×	×	×	×	×	×	×
Game gesture	×	×	×	×	×	×	×	×	
Big science	×	×	×	×	×	×	×	×	×
Data bag exploration	×	×	×	×	×	×			
Risk analysis	×	×		×	×	×	×	×	×
Churn analysis	×	×	×	×	×	×		×	

Fig. 10: Extended RDBMS system for data warehouse

information at right time. To address the future requirements, Company A needs to expand its data warehouse and business intelligence operations to make most out the business value which comes from data. Company A has different sites in Canada, Chile, Peru, Argentina, South Africa and Australia, each site has its own central offices where they perform all the business operations according to their needs which at some stage needs to standardise and make it global to produce a single solution. To address large data sets, RFID Assets, social media for legal and corporate laws, so company will analyse them properly need to address some of the changes in the future (Table 2). According to Table 2 which is explaining and hovering curtains from big data analytics which can provides benefits to Company A explain in the table in many ways.

According to Fig. 10 which is discussed about how to manage the unstructured and structure data, social media, news and blogs from anywhere from the internet. It is also explained about the integration of the data to address these requirements before entering into the data warehouse environment. Company A needs an integrated architecture which will support both traditional data warehouse and big data implementation. For the best result big data needs to be added into the existing data warehouse platform, so analysis can be used for the business intelligence system for better visualisation. Company A is using heavily Oracle products, even the hardware Exadata Solaris boxes which will make implementation and integration of big data into existing data warehouse system easier if Company A will keep the same vendor.
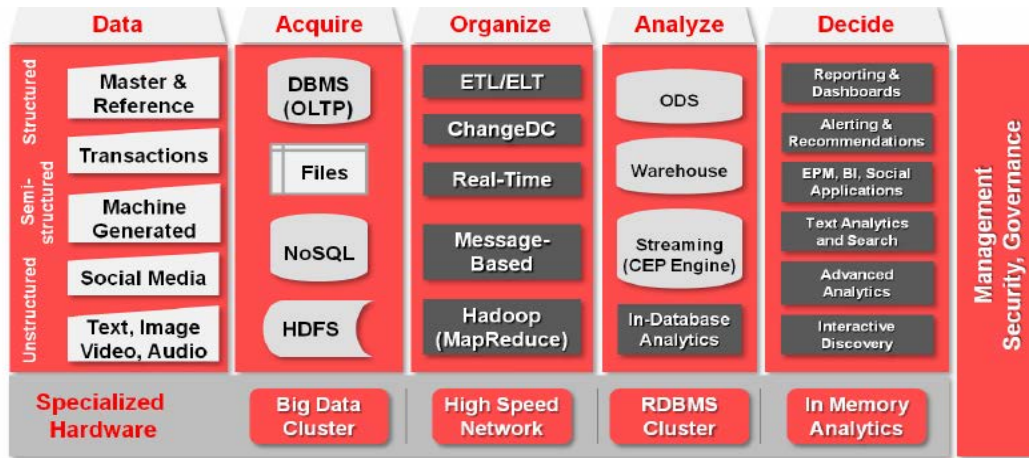
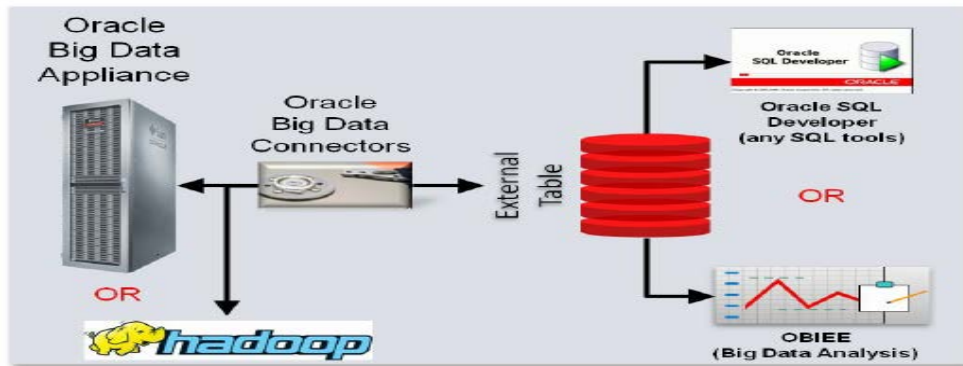Fig. 11: Oracle integrated information architect

Fig. 12: Architect decision

According to Fig. 11 which shows different data sources are captured and stored which later processed into traditional DBMS (OLTP), files, distributed clustered systems such as NoSQL and Hadoop Distributed File System (HDFS).

According to Fig. 12, this explains some of the architecture from big data appliances to the reporting level. Company A is already using Oracle stream for warehousing and reporting purposes, the only piece is missing to bring Big data analysis and tool in place for detailed level of analysis.

According to Fig. 13 which shows different data sources are captured and stored on big data appliances, this unstructured information will process into the format by using Hadoop, MapReduce and NoSQL which can make sense for RDBMS system to integrate Data Warehouse system in place. This information later processed for reporting and visualisation purposes.

Sandboxes: Kimball discussed about sandboxes which are the copies or segment of huge datasets. For the experimentation purposes Company A needs sandboxes to store segment of massive Company A datasets from different data sources. It will provide the data analyst to play with the data and do the analysis with any tool they want to use to make sense of the data to the business. The data stored on the sandboxes store for certain period of time then discard the data after the analysis.

Company A corporate private cloud: Private cloud deployed on Company A's own site or datacentre belongs to the service provider. It will give company the option where small number of business users are located provide application and data access through clouds rather than spending and investing a small data centre for 50-100 users.

Moving big data into clouds: Private cloud implementation will also give Company A the benefit to move Big data scenarios into the cloud in future if it has too. Kisker explained if information is not located inside the
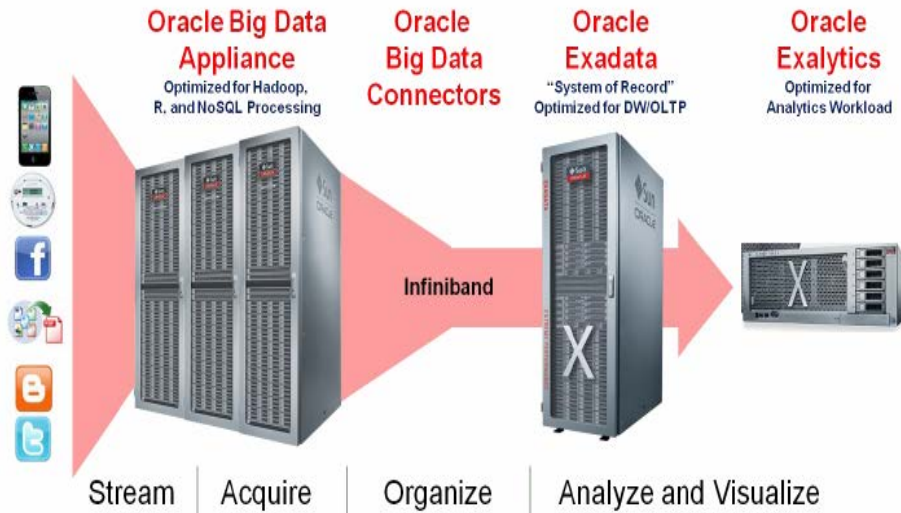
Fig. 13: Physical architect for combine analytics

Table 3: Issues identification and its solutions after implementation of big data and private cloud

Area	Present	Future with big data and clouds
Big data	Poorly manage and misplaced data; Absence of shared procedure; Missing standard approach for data standards; Unable to access old and lost information or if the information is there but not shared; Proper supervision of hardware medias and backups devices Outsourcing for backup and policies (Potential threat); Complex Scalability in existing structure; Unstructured data is not supported; Due to lack of any connectivity for unstructured data and social media, Company A was less respondent to the cases like Led Factor in children in Mt ISA, which is still in court; Lack of data quality and analysis on the datasets; Not proper data visualizations at high and global level to make decision; Missing Data Connectivity through Oracle Warehouse Builder; Additional soft wares required to connect through non-oracle data sources; Inconstant data from multiple sources will bring data discrepancies and wrong decisions making, Data is getting bigger than ever before and doing analysis on large and massive is not that easy; To manage large datasets and analysis on massive data is time consuming for financial transactions, which are getting bigger everyday; Mobile Assets and Assets issues can be resolved through big data; Mobile Assets and Assets management is the massive issue; RFID tracking is available on consignments and mine trucks but no information is stored and do analysis of RFID data; No real time analysis on the business operations, which needs advance preparations for month end processes; Frauds in Finance are paying heavy cost adding into Company A's budget value, which sometime requires PWC to find gaps between financial numbers;	Managed and low level detailed information; Data Analysis gathering requirements from process owners; Standard approach to follow data standards and business processes; HDFS will provide standard and proper backup and archival procedure which will replicate datasets, there will be no additional investment will require on backups anymore; Scalability and availability of information; Migration data from different sources, structured and unstructured data to single data warehouse through HDFS and MapReduce; NoSQL will be used to gathered all non-relevant DBMS information; Convert all the unstructured, media, web data information into RDBMS. Information will be available all time for any sources and can avoid claim cases Big data will provide more quality and detailed analysis for untouched data; Through big data advance analysis and global visualisation for business owners not only for single site, which will also provide advance statistical and predictive modelling of the data; Provide the detailed analysis for planning, budgeting and forecasting for finance subject area for better decision making; Need Oracle Data Integrator for Big data connectivity adaptor, so all the information can easily integrate in the traditional warehouse system ODI will also provide the data linages from Reporting tool to the source level; High level of data accuracy ,transparency, accountability and consistency; To handle bigger and larger data sets big data appliance and tools needs to use to make more sense of data and detailed level if information; Handling large data sets is much faster and to make right decision at right time; Provides advance analysis, data mining and prognostic modelling techniques data discovery and explore different data patterns and associations To process and analysis on RFID data is done only though Big data which will support transport and logistics; Big data will provide you to do the analysis on real time data rather than on historical data; Big data will provide the option which decision or actions have associated risk involved with it; Also provide detailed analysis for the finance data to provide fraud management and facilities the business with data inconsistency to detect the frauds
Clouds	Expensive maintaining cost when small number of business users on different site Some of the sites in Peru and Chile has only 30 users which will not going to increase in next 20 years, Company A has got datacentres and application centres, which is costing maintaining and licensing cost	Private Cloud needs to be implemented for small number of business users, which will provide low maintenance cost and no licencing cost In future big data analysis from different sites can place on the clouds and used in EDW system for analysis

company and it is coming from outside the firewall then it can be stored on the cloud for analysis. This information is related to tweets, feeds, internet, news, web data and events (Table 3).
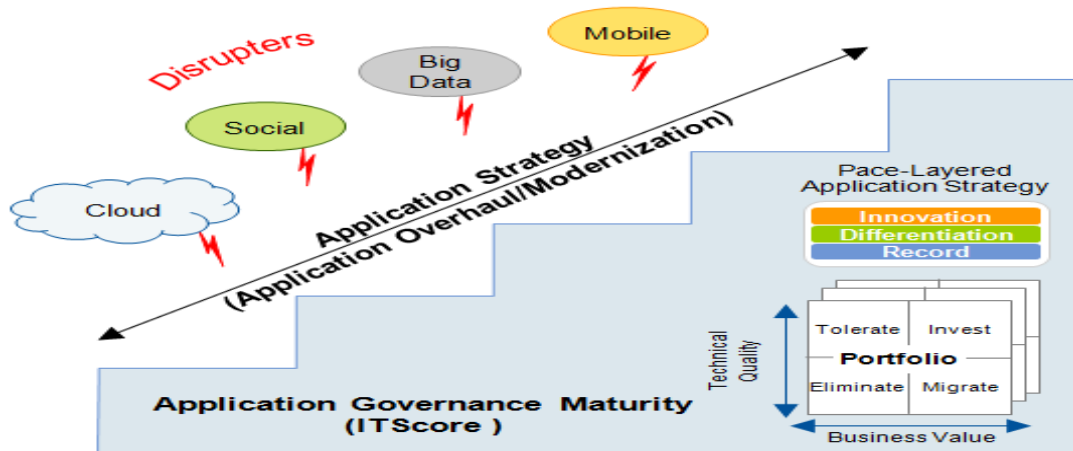
Fig. 14: Application strategy and governance agenda

Resultant ICT impact and business or regulatory requirements: Adrian and Kathy described that customer data is the precious assets for the organisations which required vigilant governance and protection. Managing big data is not easy and it will require lot of precaution and architectural framework to make sure the quality and security of the underlying data. Tankard (2012) discussed about the data governance issues related to more towards security. Big data provide the detailed level of analysis and sensitive information to trade secret, financial archives and knowledgeable belongings. Make information available centrally will provide informal and treasured target for attackers or hackers which can harm the status of the organisations.

According to Fig. 14, application policy and governance program offer organisation best practices when providing business approaches and application schemes.

Christopher discussed about cloud computing providers have extensive regularity ambiguities and absence of standardisation which needs to address that sort of issues as well. Some of the companies will provide corporate companies to do penetration testing which will highlight where the loop poles are in the implementation. Companies also implement their own security controls deployed on its own sites for cloud computing to look after the security and privacy issues. In relation to private cloud implementation for Company A requirement are given below:

- Cloud service provider (Telstra)
- Mutual agreement on contract
- Security and privacy controls form Telstra

- Security and privacy controls within company A for the security
- Penetration testing required in certain time frames which provides by different companies to find loop holes
- Applications needs to decide by the infrastructure team to place on private cloud
- Data ownership
- Make sure the regulations from the country regularity authority body if exists
- Redesign and revisited the privacy and security policy framework

David discussed about that there is lack of technical people staff with profound knowledge of analytical skills. There are some risks attached to big data such as regulatory and legal challenges which needs proper education and training to the organisations staff. Roxane and Ted discussed about data quality and governance plans need to be revisited and modified when implementing big data in organisations. Traditional data quality will not work in context of big data, organisations will be seeing the whole data sets for wholeness, correctness and reliability of the data. With big data sets, data profiling will be time consuming and inefficient, so different practices which involves limit the scope of the data to accomplish data quality processes. Some of the data flow processes happens outside the organisation's control which needs to be handle through by creating quality controls procedures otherwise there will be harmful for internal processes. In relation to big data implementation Company A will need extra hardware and applications given below:

- Big Data Appliances
- Licencing for Oracle Data Integrator (ODI), Hadoop, MapReduce and NoSQL Applications
- Transferring all the code from OWB to ODI because OWB will no longer supported and obsolete in next two years
- ODI has adaptor for big data connectivity and support heterogeneous sources
- Needs Big Data Scientists
- Staff training
- Master Data Management for Big Data
- Data Steward
- Redesign the Governance Policy Framework

CONCLUSION

Big data plays an important role in modern organizations in today's world. Organisations are managing huge volume of data, at great velocities, from multiple sources which includes structured, unstructured, conventional, non-conventional, inside and third party data and information because of big data. Complied and gathered detailed level information about customer information which was not possible before. It provides organisation ability to classify and handle all transactions, their hierarchies and association through variety of views. Big data is leading edge for the business growth and cost reduction for the organisations and delivers new openings for the CIO to become the strategic partner with different business divisions. Company A needs big data technology for its strategic objectives and to gain advantage of its competitors which will give Company A more standardise approach for business operation good data governance, detailed data analysis and information for improved decision making, fraud detection in finance and accounts subject areas, mobile assets management, fixed plant equipment, statistical analysis for correlation, regression and business trends, superior visualisation of the data and towards enhanced single approach for global KPI's and business solutions. EMC explained that big data will bring new functioning competences, fresh profits streams, modest benefits and new business model and processes. Now it is the time for the business frontrunners to reflect tactically to prepare their organisations for the big data because technology wise it is mature now and lot of well renowned business names already deployed it successfully.

With all these benefits there are some challenges when it comes to big data and cloud computing. Big data contains data ownership, data quality, data governance and lack of staff type of challenges which require proper consideration, investigation and deployment as part of big data implementation. On the other hand, in cloud computing also contains security, privacy and data ownership types of obstacles which also need proper analysis and deployment for cloud implementation. One of the main obstacle facing in the implementation of big data and clouds are governance and privacy issues which involves re-designing organisation's policies, impact analysis and mitigation strategies for the effective execution. Massive research, developments and investments are still improving the big data technology by different IT vendors which can predict the long lasting future of big data and cloud technologies in IT and business world.

REFERENCES

Agrawal, D, P. Bernstein, E. Bertino, S. Davidson and U. Dayal, 2011. Challenges and opportunities with big data: Big opportunities to create business value 2012. EMC Insurance Group Company, Des Moines, Iowa. http://www.emc.com/microsites/cio/articles/big-data-big-opportunities/LCIA-BigData-Opportunities-Value.pdf

Agrawal, D., P. Bernstein, E. Bertino, S. Davidson and U. Dayal *et al.*, 2011a. Challenges and opportunities with big data 2011-1. Purdue University, West Lafayette, Indiana.

Agrawal, D., P. Bernstein, E. Bertino, S. Davidson and U. Dayal *et al.*, 2011b. Challenges and opportunities with big data 2011-1. Purdue University, West Lafayette, Indiana.

Arnold, S.E., 2012. Big data cows and cadastres. KM. World, 1: 1-22.

Crump, G., 2012. How to protect a big data archive. Inf. Week, 133: 17-17.

Fox, B., 2012. Using big data for big impact. Health Manage. Technol., 33: 32-32.

Ghosh, D., 2010. Multiparadigm data storage for enterprise applications. IEEE. Software, 27: 57-60.

Gregory, A. and K. Hunter, 2011. Data governance-protecting and unleashing the value of your customer data assetsStage 2: Fixing the priorities; Increase revenues and profits by raising your data quality IQ. J. Direct Data Digital Marketing Pract., 13: 40-56.

Hopkins, B., 2011. Big data: How enterprise architects should exploit future opportunities. Comput. Weekly, 1: 11-11.

Janet, B. and J. Bradley, 2011. The ultimate flexible enterprise-level solution. Walmart Retail Company, Bentonville, Arkansas, http://research.microsoft.com/en-us/um/redmond/events/cloudfutures2011/slides/friday/applications3_bailey.pdf.

Matelski, J., 2012. Big data-challenge or opportunity?. Database Trends Appl., 26: 32-32.

Nasar, M. and J.V. Bomers, 2012. Data management and financial regulations: Using a big data approach to regulatory compliance. Bus. Intell. J., 17: 34-40.

Smith, M., C. Szongott, B. Henne and V.G. Voigt, 2012. Big data privacy issues in public social media. Proceedings of the 2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST), June 18-20, 2012, IEEE, Hannover, Germany, ISBN:978-1-4673-1702-3, pp: 1-6.

Tankard, C., 2012. Big data security. Network Secur., 2012: 5-8.

Valle, L.S., E. Lesser, R. Shockley, M.S. Hopkins and N. Kruschwitz, 2011. Big data analytics and the path from insights to value. MIT. Sloan Manage. Rev., 52: 21-32.

Watts, M., 2012. Capturing big datas big payoff. Money Manage. Executive, 20: 10-10.

Yoo, C.S., 2011. Cloud computing: Architectural and policy implications. Rev. Ind. Organiz., 38: 405-421.