Big Data Security

Good Practices and Recommendations on the Security of Big Data Systems

DECEMBER 2015

About ENISA

The European Union Agency for Network and Information Security (ENISA) is a centre of network and information security expertise for the EU, its member states, the private sector and Europe's citizens. ENISA works with these groups to develop advice and recommendations on good practice in information security. It assists EU member states in implementing relevant EU legislation and works to improve the resilience of Europe's critical information infrastructure and networks. ENISA seeks to enhance existing expertise in EU member states by supporting the development of cross-border communities committed to improving network and information security throughout the EU. More information about ENISA and its work can be found at www.enisa.europa.eu.

Authors

Rossen Naydenov, Dimitra Liveri, Lionel Dupre, Eftychia Chalvatzi, Christina Skouloudi.

Contact

For contacting the authors please use resilience@enisa.europa.eu
For media enquires about this paper, please use press@enisa.europa.eu

Acknowledgements

This work has been done in collaboration with CNIT (National Consortium for Telecommunications a non-profit organization of 37 Italian universities), and in particular with the experts Maria Cristina Brugnoli, Emiliano Casalicchio and Federico Morabito.

We would also like to thank the following people who helped us (in no particular order): Dragan Milisavljevic (Telecom, Austria), Gerasimos Moschonas (Alpha Bank, Greece), Annabelle Lee (Electric Power Research Institute, United States), Richard Benjamins (Telefonica, Spain), Alexandre Gaspar (Telefonica, Spain).

ISBN 978-92-9204-142-7, DOI 10.2824/13094

Table of Contents

Executive Summary 4
1. Introduction 6
1.1 Policy Context 6
1.2 Scope and objectives 6
1.3 Target audience 7
1.4 Methodology 7
1.5 Structure 7
2. Understanding Big Data 8
2.1 Overview of Big Data 8
2.2 Big Data Application Domains 9
3. Security Challenges in Big Data 13
3.1 Use case 1: Big Data in the Finance Sector 14
3.2 Use case 2: Big Data in Power utility 15
3.3 Use case 3: Big Data analytics for Telecom operator 16
3.4 List of Security Challenges 18
4. Mitigation measures and good practices 20
5. Recommendations 26
References 27

Executive Summary

We witness a new industrial revolution driven by digital data, computation and automation. Human activities, industrial processes and research, lead to a data collection and processing on an unprecedented scale, spurring new products and services as well as new business processes and scientific methodologies.[1]

Big Data is increasingly being used in many sectors. Research institutions, Industry and Government agencies active in Big Data technologies, have been working more than ever before on building novel data analysis techniques for Big Data. Business players and technology providers work on creating new products and services and even developing entirely new business models that are massively based on aggregation and analysis, of extremely large and fast growing volumes of data. It is important to understand that although most of the organizations see the potential of Big Data, they are still in the research phase, and it is very few that are actively exploiting the benefits of the technologies.

One of the main issues in using Big Data systems is security. Big Data systems are complex and heterogeneous, and the security of the whole system must be holistically approached. Moreover, the integration of different technologies introduces new security issues that must be properly addressed.

In creating this report we analysed input from a number of different sources, to better understand the usage of Big Data systems in different sectors. Based on the Big Data analysis, recommendations to organizations are provided in the current report on how to support secure adoption of Big Data systems.

In our study we have identified the following key challenges related to the secure use of Big Data:

- Access control and authentication – in a Big Data environment access to data is given to different people and entities in order to make computation and decisions. Maintaining the desired level of access control and authentication is a potential problem, as some of the entities might not have the capabilities to use the required security level.
- Secure data management – the vast amount of logs the Big Data system collects is a key issue because the big volume of logs needs to be stored and protected. Proper protection is one issue, but there is also another – it should be possible to restore them in a secure way.
- Source validation and filtering – the essential use of a Big Data system is that it could collect information from many sources. Some of these sources may not be verified or trusted. Making analysis or decisions based on input that has not been verified could lead to potentially incorrect results.

From the analysis of our interviews and surveys we have identified the following recommendations:

- Policy makers should focus on providing guidance for secure use of Big Data systems in the critical sectors.
- The standardisation bodies should adapt existing standards or create new security standards to include Big Data.
- Big Data providers or vendors should invest in compliance with security standards for their products (devices, services, cloud etc).

1 COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, Towards a thriving data-driven economy, {SWD(2014) 214 final}

- The competent authorities of the critical sectors should encourage vendors to offer security authentication mechanisms.
- Industry players and vendors should invest more into enhancing technical security skills of the staff on Big Data, through trainings and certifications.

This study presents challenges and issues identified through real life use cases and surveys. In our analysis we have identified that Telecom and research sectors seem to be the leading in Big Data utilization - something that might be due to Cloud adoption rate in these sectors.

1. Introduction

The explosion of digital technologies in recent years has enabled many opportunities which allow people and software to make real-time adjustments and decisions. Big Data is introducing the capability to sense the needs of the external market and to analyze them in a speedier and broader manner . Thus, the use of Big Data provides a competitive advantage to businesses who may identify opportunities, for executing targeted changes, to continuously improve.

According to Enisa's Threat Landscape 2014 [2], Big Data is considered as a valuable asset and as such is being targeted by cyber-attacks. Moreover, it is expected to become a very powerful tool for security professionals, as it contributes significantly to building intelligence about threats and incident management. Enisa continues to enhance the cyber security capability of the EU and its' Member States, through identifying Big Data security challenges and good practices.

In this report, Big Data is defined as the technologies, the set of tools, the data and the analytics used in processing large amount of data. While the focus of this report is the emerging information security challenges, it also assesses current adoption and the significance of Big Data in terms of business.

1.1 Policy Context

It is evident that data has become a key asset for the economy and our societies, similar to classic human and financial resources. However, Europe is lagging behind in the global market; a mere two of the top twenty companies, which use Big Data in a significant way, are in the European Union. This situation needs to experience a turnaround. The European Commission responded to the needs to strengthen all aspects of the "data value chain", to enable the evolution of a dynamic Big Data value synthesis, by conducting studies and initiatives - such as "Towards a thriving data-driven economy"[3] and the "Worldwide Big Data Technology and Services - 2012-2015 Forecast"[4]. Additionally, the European Commission has formed the Big Data Value Public Private Partnership (PPP), to cooperate in data-related research and innovation, enhance community building around data, and set the grounds for a thriving data-driven economy in Europe[5].

1.2 Scope and objectives

The study aims at identifying the key security challenges that the companies are facing when implementing Big Data solutions, from infrastructures to analytics applications, and how those are mitigated. The analysis focuses on the use of Big Data by private organisations in given sectors (e.g. Finance, Energy, Telecom). However, more institutions (e.g. research centres, public organisations, and government agencies) has been considered as well.

More specifically the objectives are:

- present a systematic overview of the security challenges for Big Data infrastructures in the EU landscape,

[2] https://www.enisa.europa.eu/activities/risk-management/evolving-threat-environment/enisa-threat-landscape/enisa-threat-landscape-2014/at_download/fullReport

3 Digital Agenda in the Europe 2020 strategy: http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=6210

4 Digital Agenda in the Europe 2020 strategy : http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=6242

[5] Digital Agenda in the Europe 2020 strategy : http://ec.europa.eu/digital-agenda/en/big-data-value-public-private-partnership

- identify emerging use cases deployed in different critical sectors (by analysing pilot Big Data adopters),
- identify per use case the security and resilience challenges of the use of Big Data services,
- define recommendations for strengthening the security of Big Data infrastructures and services.

1.3 Target audience

The document addresses Big Data adopters as well as institutions in Europe evaluating the use of Big Data services. The work has been conducted in order to provide findings, good practices, guidelines and recommendations suitable for a number of different sectors (Health, Transport, Energy, Finance, etc.). The document aims to serve the personnel responsible for IT and/or innovation activities in their organisation, and in particular:

- IT/Security Solution Architects;
- CISOs;
- Information Security Experts
- CEO, Directors
- Project Managers
- Decision makers in general

1.4 Methodology

This study and its outcome are based on stock taking, and using different data collection techniques:

- State of the art study and desktop research, during which we investigated the European landscape, the purpose and the services Big Data cover, as well as the security requirements and controls in place;
- Online survey and in depth interviews with experts from private and public institutions implementing or using Big Data services, providers of Big Data services for commercial purpose, experts studying Big Data for investigation purposes and stakeholders;
- Analysis of the findings and definition of Use Cases.

1.5 Structure

The report is organized in 4 main sections.

Section 1 introduces the objectives of the study, the target audience, the methodology used to collect and analyse data.

Section 2 sets the background for the study. First we discuss the definition of Big Data from the literature. Second, we present and define the Big Data ecosystem - illustrating the Reference Architecture for Big Data systems. Finally, we report about the European landscape on Big Data usage.

Section 3 presents the security challenges resulted from the survey and the interviews conducted, it analyses the data collected through three use cases which represent the main security issues and good practices adopted.

Section 4 presents the mitigation measures and identified in the previous section providing sectorial, as well as horizontal solutions.

Section 5 draws the final outcomes and conclusions by the good security practices emerged from the use cases. Finally we give five recommendations and conclusions to implement secure diffusion of Big Data in the European landscape.

2. Understanding Big Data

2.1 Overview of Big Data

An increasing number of devices, sensors and people are connected to the global network and this changes dramatically the ability to generate, communicate, share and access data. Therefore, the data volume has become so large in recent years, that it cannot be processed using conventional methods.

In the past, the rapid creation of such volume and variety of data would have caused significant problems. Nowadays, with decreasing storage costs, better storage solutions and algorithms to create meaning from all that data, this is considered rather an opportunity, than a problem.

In 2014, EMC/IDC[6] extended the definitions describing Big Data technologies "as a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-Velocity capture, discovery, and/or analysis.

Data Volume means the size of data, data Velocity the speed at which new data arrives and Variety means, that data is extracted from varied sources and can be unstructured or semi structured. Actually, in industry Veracity is sometimes used as a 4th 'V', referring to the trust into the data and three other V's are becoming used for the definition of Big Data: Value refers to the inherent wealth, economic and social, embedded in any data set; Volatility refers to the tendency for data structures to change over time; Validity refers to appropriateness of the data for its intended use.

There are three main parts of Big Data system: the data itself, the analytics of the data, and the presentation of the results of the analytics. Big enterprises and industries, are primarily focusing on Big Data analytics to improve their operations in terms of marketing, supply chain, information security etc. Industries get a huge benefit from analytics-driven insight that enable them to track and manage operations in a more intelligent manner. Many companies have dramatically boosted profits and have met consumer demands more proactively, by utilizing automated data collection to feed information into a big data analytics program. Big Data include nine characteristics which assist in recognising it when some or most of them are in place.

- **Fast data insertion.** Vast amounts of data generated every second are stored and analysed.
- **Distributed redundant data storage.** In Big Data storage method is based on a distributed file system that gives the needed redundancy and high availability.
- **Parallel task processing.** Computation is performed in parallel and large volumes of unstructured data can be efficiently processed in a few minutes.
- **Different types of data.** Big Data technology enables concetration and analysis of unstractured data, such as conversations, videos, images, sensor data, etc.
- **Scalable.** Big Data systems store and distribute very large data sets across a vast number of systems that operate in parallel.
- **Large scale analytics.** Fast data insertion – as mentioned above - creates an enormous amount of data, which is then analysed to produce large scale analytics which contribute to a better planning or management of the area they fit (fit-of-purpose governance).

[6] EMC/IDC, "The Digital Universe" Study, 2014

- **Hardware agnostic.** Big Data processing and Big Data analytics is executed efficiently regardless the underlying infrastructure, resulting in an improved direction and decision making around hardware investments.
- **Accessible.** Easily access new data sources and tap into different types of data (both structured and unstructured) to generate value from that data.
- **Cost effective.** Big Data systems also tackle the problem of traditional relational database management systems (RDBMS). Use of databases specifically engineered for managing large volumes of data, where traditional RDBMS will be extremely expensive.

2.2 Big Data Application Domains

Private companies and the public sector in the EU have started considering Big Data solutions to add value to their business services and to optimize their internal processes.

Most of the potential adopters are currently in the business requirements collection phase. Hence, architectures and platforms that would form Big Data analysis systems have not been defined yet, or are in very early stages. Furthermore, there are a few examples of Big Data applications, which are either in the design or implementation phase. These examples assist to clarify trends and challenges of Big Data, that need to be addressed to accelerate the deployment of secure Big Data solutions in the European landscape.

To provide a general representation of the Big Data application purposes ecosystem, we provide a mapping of typical high level application, purposed per sector in figure 1, having identified some essential processes and services that are based on Big Data.

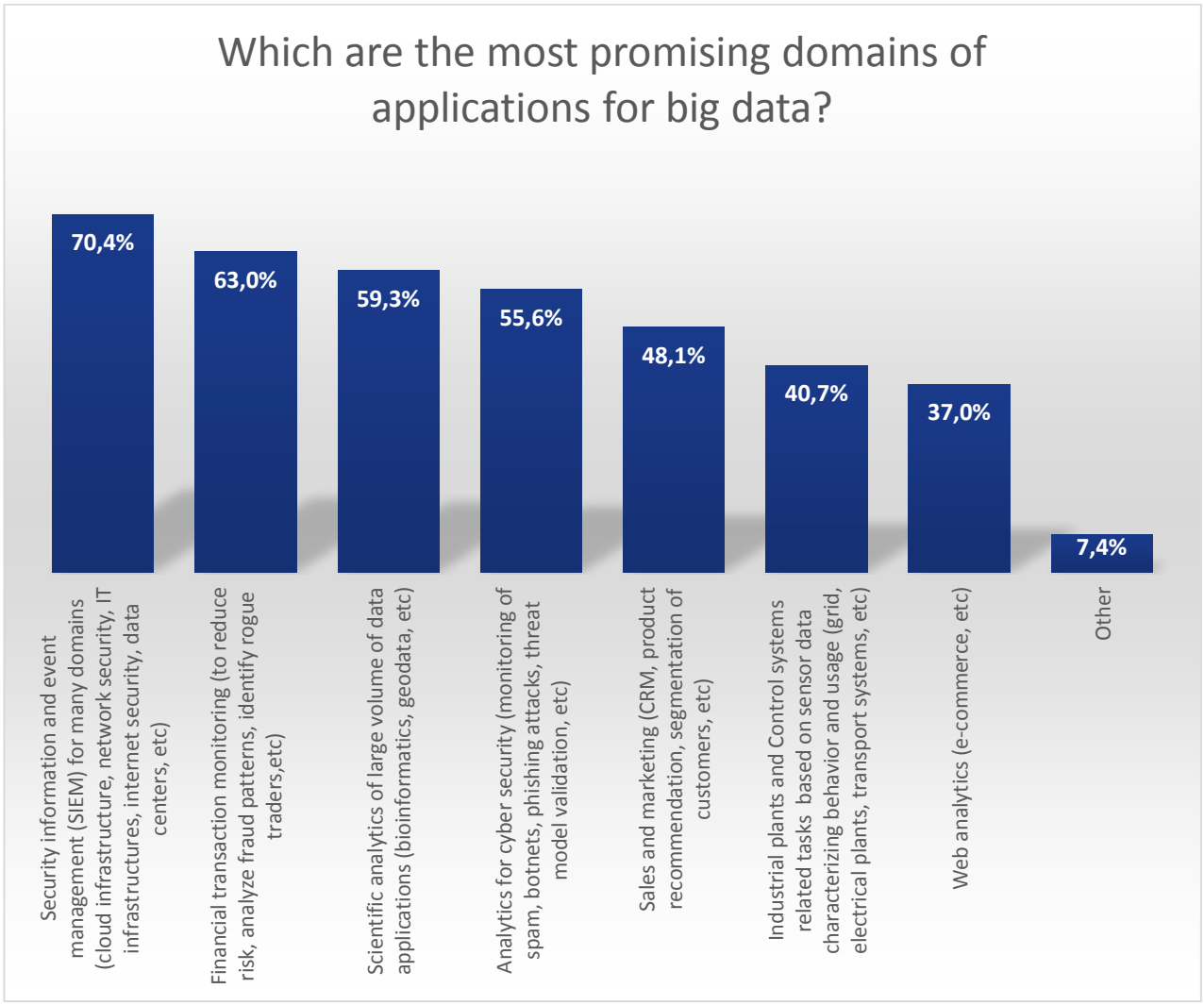

Fig. 1 "Which are the most promising domains of applications for Big Data?"

Application purposes of Big Data are multiple: although in the study a significant number of cases have been collected, we provide some examples that do not cover the entire range of Big Data applications.

In particular we have considered the following main examples:

Advanced security information and event management (SIEM): Nowadays, in the Big Data era, second generation SIEM tools are able to capture unstructured data, that is becoming relevant to enterprise security from all over an organisation and carry out complex queries and receive results in a timely fashion. This data, which stems from modern security priorities such as web and email, is hard to analyse and make sense of using SIEMs alone. Unstructured data was difficult to capture by utilizing first-gen SIEM technologies. One of the advantages of Big Data and NoSQL databases, is that they can store such data in a format that is scalable and at the same time allow to

create queries and understand the data better. In reality Big Data offers as added value not only the correlation between unstructured data, but also SIEM scalability.[7][8]

Financial transaction monitoring: The finance sector is an industry that deals with a lot of fraud. The benefits that the FIs are seeing are the possibilities to combine different sources of data, and perform analytics in real-time to identify fraudulent actions. Multiple insurance companies are also seeing the benefit of it. Similar methods of operation could potentially be quickly recognized and prevent fraudulent claims. The FIs are able to analyze an enormous amount of data from unstructured sources – such as social networks are - and fuse it with their internal data to take quick decisions.

Scientific analytics: Many research and governmental institutions are exploring innovative approaches for social security. Examples of such possibilities are homeland security through analysis of social networks, and financial transactions of possible terrorists, national security through the management of extremely large video and images datasets collected by the satellites. At the same time, the result of the analysis of satellite data can be sensitive and then classified and security requirements of the filtering and processing solutions of the Big Data services should be clearly defined. In the field of astronomy, Object catalogues, imaging data, and spectra have been publicly released through the Sloan Digital Sky Survey web site[9], along with detailed documentation and powerful search tools. Practices within the scientific community are starting with the aim of publishing data to open access databases and repositories, which could then be reused for a variety of different purposes. In Health-care, examples of the areas of Big Data applications are medical imaging archiving and processing, personal health records archiving, processing and sharing, analytics for unified electronic health records (EHRs) and EHRs radiology images and genomic data. In Transport sector the applications of Big Data are multiple - visualization of traffic and resources, optimizing the transport services and maximizing the availability of the transport resources.

Cyber Security analytics: Big Data is changing the security landscape, and what IT professionals need to know to stay abreast of the new approaches as it enables various capabilities; for instance, the analysis of long-term historical trends and predictive analysis**.** By collecting data on a large scale and analyzing historical trends, it is possible to identify when an attack started, and what were the steps that the attacker took to get ahold of your systems. Even if they did not detect the original attack, they can go back to carry out a historical correlation in the database and systems to identify the attack. Additionally, the efficiency of queries allows the transformation of collected data to meaningful information. Big Data enables IT professionals to carry out complex queries and receive results quicker than in any other case. Security vendors are providing new threat-reputation services and analytics tools, which introduce automated approaches and high-performance security monitoring systems that keep up with complex real-time analytics to meet detection and performance requirements. Big Data analytics can track trends and security breaches, and allow companies to proactively go after threats before they strike (predictive analysis). Rabobank[10] is a bank that has tested a system that analysed criminal activities at ATMs, to determine factors (i.e. proximity to highways, weather condition and seasons) that increase the risk to becoming victimized.

Sales and marketing: Customer Relationship Management fits well with the Big Data technologies. Having the possibility to enhance the knowledge about customers from sources other then the traditional, gives the organizations opportunities for better sales, customer needs and segmentation analysis. In Telecoms, Big Data is perceived as a strategic opportunity for the development of innovative services, that exploit the unique potential

[7] http://www.information-age.com/technology/security/123458055/big-security-big-data-and-end-siem#sthash.o5uPQvbi.dpuf
[8] https://cloudsecurityalliance.org/download/big-data-analytics-for-security-intelligence/
[9] http://www.sdss.org/
[10] http://www.emerce.nl/nieuws/big-data-nieuwe-olie

for the amount and quality of data collected through their networks. Telecoms use multiple indicators, such as billing and sentiment analysis, to identify customers that can be upgraded to other products, as well as to select those to high lifetime customer-value, so its' team can focus on retaining those. FIs use data analytics to predict which financial products and services would a customer appreciate, so they can better target consumers during sales process with these insights. In Heineken,[11] one project provides real-time personalized marketing messages to fans who happen to be watching a sponsored event. Spotify[12] uses data from user profiles, playlist and historical data to provide recommendations for each user.

Industrial control systems: In the Energy sector for example Big Data and analytics are one of the new enablers for improving the visibility of the Smart grids, and give the utilities operators a new way for the planning, operations, and maintenance of their power system. CERN[13] is using Big Data to identify abnormal behaviour, allowing a better system maintenance through identifying failure patterns and automatically executing correcting actions.

Web analytics: In our present time everything is moved to the web. Being able to identify how customers behave after visiting a web site, and what are the differences could potentially allow for great improvement. Web analytics are a key component for realising better product development and better customer experience. In web analytics it is not just web site data that is being used. Advertisement, market data, customer profile data and social media data are just some of the sources that may be taken into account when doing web analytics. This requires a system that can collect and analyse such volume of information and produce results, in some cases even in realtime.

[11] https://datafloq.com/read/how-heineken-interacts-with-customers-using-big-da/384
[12] https://labs.spotify.com/2013/05/13/analytics-at-spotify/
[13] http://openlab.web.cern.ch/sites/openlab.web.cern.ch/files/technical_documents/Data.Analytics.on_.Control.Data-POSTER-_BoS_2014.pdf

3. Security Challenges in Big Data

Securing Big Data comes with unique challenges, beyond being a high-value and attractive target. It is not that Big Data security is fundamentally different from traditional data security. Big Data security challenges arise because of additional differences, not fundamental ones. The differences between Big Data environments and traditional data environments include:

- The data collected, aggregated, and analyzed for big data analysis
- The infrastructure used to store and house big data
- The technologies applied to analyze structured and unstructured big data

Since the main priority is to offer speed for a great volume of data, often security can be the last item to consider; mainly because there is no specific classification of the data that will be stored and transferred. The integration of different technologies introduces new security challenges that need to be properly addressed, by usually broken down into technology specific challenges. In case that Big Data systems are supporting critical infrastructures security becomes a requirement as well. Since Big Data systems are complex and heterogeneous, the security approach must be holistic in order to ensure availability and continuity of the services.

To understand the security challenges, we decided to follow examples of Big Data usage and study the specific cases that apply. This section analyses three (3) relevant case studies, and describes the idendified security challenges. The criteria for selecting these use cases is the maturity of Big Data in these sectors, providing greater value from the collected information. It must be taken into account that since Big Data is a relatively new notion, the community is in its early stages in identifying solutions on the specific issues, such as security. We depict below the impact based analysis for these use cases:

USE CASE	BIG DATA APPLICATION	CRITICALITY
Financial Sector	User of Big Data	Integrity of data and information
Power Supply (energy sector)	User of Big Data	Availability and continuity of service
Telecommunications Sector	Provider of Big Data	Secure provision of services

Table 1 Criticality of Information Security for Use Cases

3.1 Use case 1: Big Data in the Finance Sector

TITLE	DESCRIPTION
Sector	Finance Sector
Usage of Big Data	Analytics, data driven decision making, real time services offering, risk quantification and prediction models building, fraud patterns analysis, rogue users identification
Security Challenges	• Trustworthiness of devices collecting data • Source validation and filtering of data • Application software security • Access control and authentication • Interoperability of devices • Distributed systems security (DDoS attack)

Table 2 Finance sector use case

One of the standard tasks the Chief Information (Security) Officer (CIO or CISO) of a financial institution (FI) has, is to use effectively the large amount of data generated every day for the support and operations of the business - taking into account the high security requirements. The CISO has to define appropriate security policies (e.g. trigger setting and device configuration) and take other security measures, as well as to show compliance with security standards to regulators and auditors. The incremental security requirements[14] lead to new approaches for using Big Data capabilities, to monitor and analyse information of the IT infrastructure.

Deployment of Big Data solutions in banks[15] for fraud detection, instead of using the Security Information and Event Management (SIEM) system is very popular. The system collects information on security events (software application events, anti-virus applications) and logs from devices (firewall, database, network events) of the IT infrastructure of the bank and its subsidiaries, handling in real time the security risks and targeted attacks to its systems' infrastructures. Big Data analysis can help detecting threats at early stage through combinatory computations, which is often impossible through traditional SIEM systems.

The application provider that monitors data transferring is a third party security company that analyses the stream data both off-line and real-time, providing feedback to the FI. The output of the analysis is based on a security model that can generate alerts in case of incidents or suspicious analysis, by setting and adjusting triggers with a level of severity depending on the criticality of assets. Security operations team will then be able to reconfigure systems in a proper way, and respond faster to security risks, activating countermeasures, or enable the security systems to automatically mitigate the threats in the future.

Security challenges in Big Data SIEM in a Financial Institution

Trustworthiness of devices collecting data: When the SIEM application would extend scope and include information and event collection from other sources (like users behaviour through active directory and email logs, bank

[14] http://www.computerweekly.com/feature/How-the-financial-services-sector-uses-big-data-analytics-to-predict-client-behaviour
[15] http://www.computerweekly.com/feature/Information-security-is-a-big-data-issue

transactions and social media) the challenge would be to validate and filter that information. In a Big Data system, data is collected from many sites sometimes of unknown credibility. Even if the Finance sector is a highly regulated one, data can be collected from non-trustworthy sources, mostly to conduct behaviour and predictive analytics.

Source validation and data filtering: Since many of the sources may not be trusted, the information collected could be compromised (integrity of data risk). The advanced SIEM for the FI described above collects event logs from a diverse number of hardware devices, software applications in the bank enterprise network, as well as social media. A key security challenge in this type of data collection process, is source validation and how can the Big Data application trust such data. In order to do that, the processing application must validate that a source of input data is trusted and not malicious. Also there are a number of filtering rules that apply to input from the data sources collection.

Application software security: Big Data implementations typically use software that was designed to manage large sets of data, and security was not priority. This potentially could lead to security issues as the software is not tested for common vulnerabilities such as back doors, default credentials and weak or no authentication methods. In addition to this, relational databases include some security features, but to use these features, data needs to be classified; in the case of other types of respositories, security is weakened in favour of flexibility and speed.

Access control and authentication: User authentication and access to data from multiple locations may not be sufficiently controlled. This is a great challenge the operator has to implement; however, the direct impact is with the user - especially in cases where confidentiality and integrity of data are a priority.

Interoperability of applications: Systems' integration can result in security gaps that emerge due to the existing differences between the platforms. This might not be a cyber security challenge per se, but vulnerabilities increase the attack surface of an adversary. This would apply to all distributed systems' settings. However, because of the great number of applications interrelating in the Big Data model, this security challenge becomes a priority. Integration and correlation of different components like workstations, application servers, databases, and network devices can be a complex problem. Interoperability between devices, due to different formats, syntax and platforms, makes communication even more difficult in processing and analysis.

3.2 Use case 2: Big Data in Power utility

TITLE	DESCRIPTION
Sector	Energy Sector
Usage of Big Data	Analytics, data driven decision making, risk quantification and prediction models building
Security Challenges	• Source validation and filtering of data • Application software security • Infrastructure Security • Distributed systems security (DDoS attack) • Access control and authentication

Table 3 Energy sector use case

In the utility use case, the use of Big Data serves the management of power infrastructure. The Big Data system aggregates and analyses data from smart meters and sensors distributed over the power grid nodes. The goals of the Big Data analysis is to develop a scalable application for management of the infrastructure, and customer satisfaction.

The Big Data system collects and integrates measures of the power assets with traditional physical modules for power plants, usinga mathematical model to identify and categorize faults and causes of the electric grid. This vast amount of data produced by smart meters, enable benefits through analytics that were never possible before. The management of the infrastructure through Big Data analytics will provide added-value services, like increased customer satisfaction, better capital spenditure (e.g. size equipment based on better load profiles, optimized asset life cycle), improved efficiency in infrastructure management, improved reliability, and resiliency of electric grid (e.g. predict and prevent equipment failure, better manage operating parameters), and improved cost position (e.g. Improve economic dispatch based on grid conditions, condition-based maintenance).

The data sources are distributed within the power assets (power distribution sites, digital relays) and the homes of the subscribers, where most of the smart meters are. The data is then integrated with other sources that come from external interfaces (e.g. weather forecasts) and internal databases (archival on consumers, facilities, etc.), posing challenges for the application framework that integrates a variety of multiple data sets.

Security challenges for the use of Big Data for infrastructures and customer management in power utility

Source validation and filtering of data: The organisation could decide to give different levels of trust and protection of data sources. These levels are differentiated according to the level of criticality of the power devices which produce them. There are two levels of protection for the transmitted data , based on the impact that a compromise could cause on the behaviour of the electric plant. For example, digital relays are considered critical because any compromise could alter the functionality of the system. On the contrary, battery management units are of low criticality, since they have a smaller impact on the continuity of the service. The level of trust in smart meters is considered low as well, as they are out in the field exposed to the outer elements, or in the users' residences, where denial of access limits the control.

Application software security: As stated above, Big Data often use combination of open and closed source software. Having multiple software development models increases the probability of having security issues within the Big Data system.

Infrastructure security: Since smart meters are the main devices to collect information, infrastructure security is a great challenge. The notion of cyber physical security, has become another important aspect we should investigate when discussing cyber security challenges. Big data is the system mostly used for the collection and analysis of data coming from sensor networks, smart meters, etc - so this challenge is directly linked to the Big Data implementation.

Distributed DoS attack: A standard challenge in distributed systems that could have great impact on the availability of the service. Combined with the cyber physical systems in place using Big Data (due to large volume of information), this is one of the greatest threats for the Big Data system. For example if a DDoS attack is launched the analysis making process could be stalled and this could lead to wrong decision or decision not done in time.

Access control and authentication: The possibility to act automatically from certain energy switches, raises a security concern, as the switches need to have access to data to take action. Since the majority of such switches are not designed with security in mind, they may not have the possibilities to authenticate to the system when accessing data, which presents a key access control problem.

3.3 Use case 3: Big Data analytics for Telecom operator

TITLE	DESCRIPTION
Sector	Telecommunications

Usage of Big Data	Big Data provider: Increase volume for data storage, services optimisation, adaptive e-services offering, real time services offering, data analytics, prediction models offering data driven decision making, risk quantification
Security Challenges	• Source validation and filtering • Application software security • Access Control and authentication • Supply chain security • Secure data management • Infrastructure security • Secure Cloud use

Table 4 Telecoms sector use case

Telecom operators are exploiting ways to develop innovative data-centric services, that can offer added value from the analysis of high quality and quantity of data generated by their infrastructure. For the telecom operators Big Data is a business solution that goes beyond optimising their in-house processes. They have various ways to develop Big Data applications using their infrastructure, offering these services to a number of other industries, including but not limited to: retail, financial services, healthcare, and marketing. So they are not only users, they are also providers. In this use case we will present the security challenges the provider of Big Data services should take into account.

Big Data has many applications in the telecom industry. One key use is data analytics for the management of telecom infrastructure. Information is collected to identify trends and analyse the behaviour of the network. Data analytics are also useful for internal services (e.g. resource management, turnover management, procurement, etc.) that process intensively generated data and integrate Big Data techniques to optimize internal processes, operations or to market products more effectively. Additionally, with Big Data the information generated from a mobile network, specifically location data, is analysed resulting in trends and patterns to better understand consumer behaviour based on mobile events that occur continuously throughout the year. The data is collected and aggregated through dedicated interfaces from the network equipment (base stations, antennas and network controllers).

In this use case we study the security challenges from the Big Data services provider side. Enriching the mobile data collected with demographic and behaviour information of the users - such as mode of transport and social events - provides sophisticated profiling and segmentation analysis services that can result to added value for multiple sectors. One example of sectors that could use this service are the transport sector, where the ability to track the location of mobile devices can assist monitoring of current traffic to save time and reduce congestion and can suggest new location-based services. Another is retail, with solutions based on the movements of the user mobile phones to provide location for a new cash dispersing machine or new branch. Offering this type of Big Data service requires the development of solutions for the collection, transport, storage, provision and use of large sets of data, covering issues such as resiliency and availability (especially for the transport sector), open data access, data and metadata quality. The telecommunication companies can share this profile, taking into account the specific security challenges that might come with it.

Security challenges for Big Data analytics for Telecom operators

Source Validation and Filtering: The aggregation of individual data coming from the mobile network, pose relevant security challenges, and specifically for data provenance control. One problem is the Big Data application framework collecting data that is generated in the equipment of the mobile network, which is composed of multi-vendor technologies. The telecom operator should verify and trust the infrastructure components - both hardware and software - that produce the data and events collected by the Big Data application framework, and ensure the proper security provenance of the data.

Application Software Security: Use secure versions of software. As described above, Big Data technologies weren't originally designed with security in mind. Using open-source technologies (which might be a cost efficient solution) can result into security vulnerabilities that nobody invested into addressing.

Access Control and Authentication: A Big Data provider should include in the security provisions access control and authentication mechnisms for the different roles a user/customer would have when utilising the service. However, the biggest challenge lie to the fact that user authentication and data access from multiple locations, may not be sufficiently controlled, due to the decentraliseed model and the various interactions of systems. Particularly in regulated industries, securing privileged user access must be a top priority. Certain users must be permitted access to highly sensitive data in specific business processes, but avoiding potential misuse of data can be problematic. Securing privileged user access requires well-defined security policies and controls, that deploy role-based access - preventing data access when the relevant authorization does not exist.

Supply Chain Security: Due to the distributed nature of the system different devices (smart or not) from different vendors are used. This represents a supply chain security problem, as the control of the devices lies in the hands of the organizations involved in the delivery of the products and services.

Secure Data Management: In case of Big Data services provisioning, security should be in scope as part of the system. Storage security should be addressed from the requirements phase. However, this would depend on the business model followed; e.g. the services can be built in-house completely, or could be bought/rented by third party providers.

Infrastructure Security: In the case of a Big Data provider infrastructure, security does not only concern the cyber-physical systems (e.g. sensor networks), but also the terminals of the end user. Should this be a portable or non portable device, the approach is posing more challenges to be addressed. A number of untrusted devices can be connected at any time in the company network, and handle company specific data, raising the issue of end-point security.

Secure Use of Cloud computing: Big Data relies heavily on the Cloud, and as such the Cloud security risks must be addressed. Issues such as vendor lock-in, isolation failures and data governance must be addressed.

3.4 List of Security Challenges

Through our use cases and survey we have identified the following Big Data security challenges:

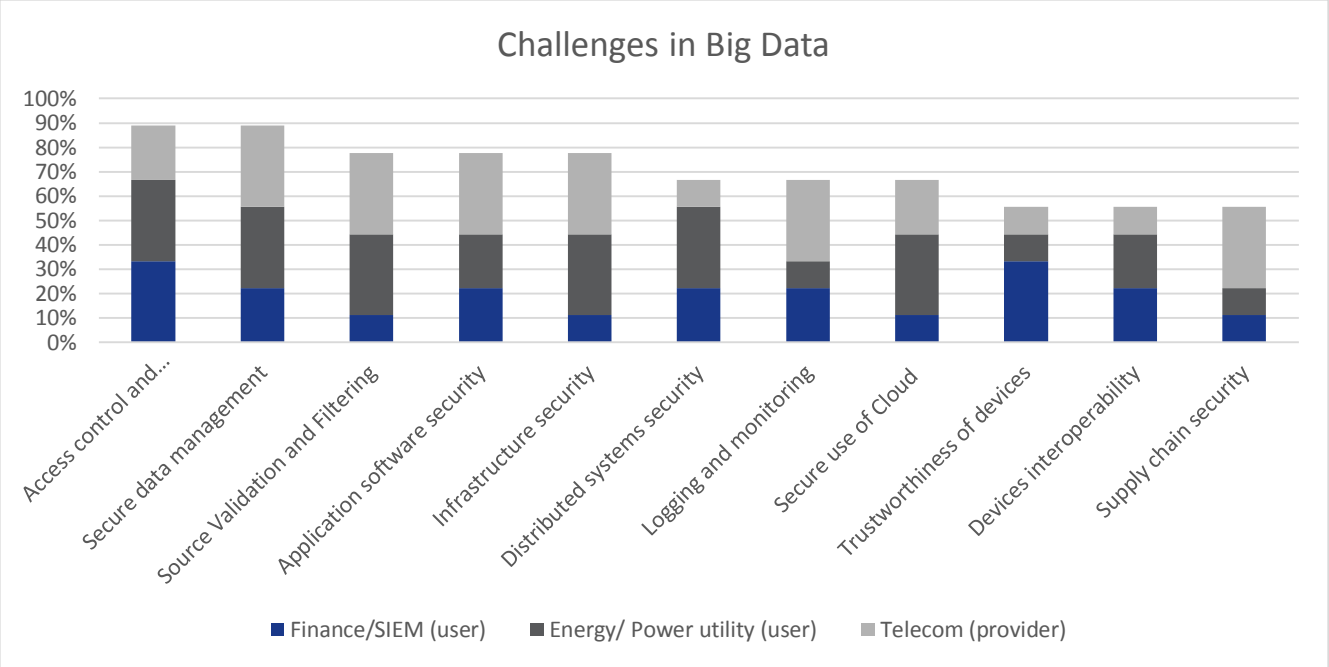

Figure 1 Big Data challenges rated by use case

Based on our analysis and the information collected the most prominent security challenges are therefore:

- Access control and authentication
- Secure data management
- Source validation and filtering
- Application software security
- Infrastructure security

4. Mitigation measures and good practices

A Big Data system is a high-value target for adversaries, since the data it contains can provide more information than simple analytics. Based on the challenges already identified, in this chapter we present the good practices and mitigation measures including examples from the use cases.

Strong and scalable encryption: Secure data management is of a great importance as it relates to both storage and destruction of data. Encryption is a key mitigation measure to ensure that only authorized people and entities have access to data.

In the Big Data ecosystem the computational capabilities could also challenge the strength of crypto algorithms. We can expect novel cryptographic or even new technologies from unchartered territories to emerge, in order to handle the need for robust and scalable encryption. New types of encryption allow for computation without the need of revealing the private key for decryption. This allows third parties to do computation with the results being known only by the owner of the private key.

In the case of the FI transactions and information from social media, the data protection and storage requirements will change according to regulations from the Financial sector. In this case (SIEM) encryption was only used during the transit phase, when delivering the information from the sources to the system. In the Big Data system itself no encryption was used. This is also true for the other two use cases (Energy and Telecom). In the latter two use cases no encryption was used for the transport of data, as the devices smart meters and base stations do not offer such a capability.

Strong and scalable encryption

- Encrypt data in transit and at rest, to ensure data confidentiality and integrity.
- Ensure proper encryption key management solution, considering the vast amount of devices to cover.
- Consider the timeframe for which the data should be kept - data protection regulation might require that you dispose of some data, due to its nature after certain period of time.
- Design databases with confidentiality in mind – for example, any confidential data could be contained in separate fields, so that they can be easily filtered out and/or encrypted.

Security testing procedures and code auditing: Big Data systems can use multiple applications for different purposes. In some cases one application is used for computation and another for visualisation of the data.

Equipment manufacturers increasingly rely on components delivered by third parties. In all our use cases we have organizations that are extensively relying on multiple vendors for their equipment and applications. In the context of Big Data, using products or applications which are not tested or certified, could lead to different problems. For instance in the Telecom use case, tampered equipment could leak processed data. The same is applicable for Energy smart meters; if their data could be snooped by an outsider, this could provide useful information for criminals to

identify whether a house is empty or not. One mitigation measure that could be used to ensure supply chain security, is to test the equipment for security weaknesses. Another mitigation measure could be to ensure in the procurement procedures that only equipment from trusted sources is used.

Application security vulnerabilities can lead to a compromise of the data integrity, resulting into incorrect findngs, even when the data is correct. Also, there are cases with two or more applications interacting with each other; where certain security measures need to be employed, to sustain the security posture. One method to follow is source code review to identify certain gaps or misberhaviour in the security of the application. Another possible mitigation is to conduct regular security tests on the applications in order to confirm the security posture.

In the finance use case, the application is commercial software. In the cases of Telecom and Energy the applications are using open source tools. However, for the Big Data provider use case (Telecom), the organisation should ensure that the applications used are secure and do not have vulnerabilities. In order to do so, the same measures could apply - like performing penetration testing, creating updating and patching procedures and enforce source code review.

Tamper-resistant devices can be used to improve Infrastucture security. Many of the devices collecting information are exposed to the elements of nature, as well as malicious actions. Some of the devices are in private properties and can only be accessed after owners' permission.

In the use case of SIEM and Energy smart meters, there were legacy devices, which do not provide the means for authentication or support for secure protocols. In these cases this funstionality was provided through the use of third party products.

Application security

- Use regular security testing procedures to re-assure the level of security, specially after patches or functionality changes.
- Ensure tamper resistant devices to avoid misuse.
- Ensure internal security testing procedures for new and updated components are carried out regularly; if it is not possible third party evaluations, audits and certification are key elements for the confidence and trust in products and actors.
- Ensure procurement policies cover purchasing from authentic suppliers.

Compliance to standards and certification could help ensure interoperability between applications, as well as ensuring that security is kept at the required level. This is also valid for connecting and filtering sources. Certification ensures that these devices can connect and authenticate without compromising security.

ENISA has created a list of certifications [16]for use in Cloud computing that could be a starting point for a similar list for adoption in Big Data.

[16] https://resilience.enisa.europa.eu/cloud-computing-certification

Standards and Certification

- Use devices which comply with desired security standards.
- Ensure obtained certification relates to the use of Big Data.

Risk assessment: Security risks associated with the use of Cloud have already been addressed in several other studies[17] from ENISA. Some of them are relate to Cloud computing exasperating the differences between different countries and different jurisdictions, sector specific regulations (e.g. Telecom, Energy, Finance) are also very different throughout EU and have a great impact on use of Cloud computing. This also relates to the issue of removing or deleting information no longer needed. In the use of Cloud computing it could take months to erase all data and this might be in contradiction with Data Protection Authorities. Another issue with Cloud is resource isolation, where one component can influence other resources. Specifically in Big Data resource influence is amplified due to the sheer volume of resources used and this can lead to wrong results or delays in the computation.

In order to ensure that Cloud could be used for Big Data systems as a starting point Service level agreements (SLA) and certifcations could be used. Certifications achieved against secure guidelines for use in Big Data systems could be a good starting point for users of Big Data.

Although many sectors now use Cloud, there is still some organizations which have not created a Cloud risk assessment. Creating Cloud risk assessment should be used to identify risks and aproprate mitigation measures.

In the use case 1 (SIEM) the system does not use cloud computing technology. In the other two use cases (Telecom and Energy smart meters) both use Cloud computing technologies for their Big Data systems.

Secure use of Cloud in Big Data

- Ensure Big Data is included in the risk assessment for Cloud.
- Ensure proper Service Level Agreements have been adopted.
- Ensure proper resource isolation and exit strategies have been negotiated

Source filtering: In a Big Data implementation, multiple endpoints may send data for processing and storage. To ensure only trusted endpoints are sending data and that false or malicious data is not accepted, organizations need to validate the authenticity of each source sending data. Host-based and mobile device security controls could potentially mitigate the risk associated with untrusted endpoints, along with strong processes around system inventory tracking and maintenance.

Another challenge that is potentially directed to Big Data systems is Distributed Denial Of Service attacks, which inherent to distributed systems. Source filtering could be based on security protocols supported and the trusted platform modules of the devices.

[17] https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing

In use case 1 (SIEM) endpoint validation is done throught the SIEM system. New sources are added only after they have been approved by the system administrator. Also sources have confidence levels and based on confidence levels actions may be triggered.

In use case 2 (Telecom) endpoints are both base stations (mobile and stationary) and mobile phones. Validation for base stations is done through inventory tracking. Data from mobile phones is not validated and is only used anonymized and in aggregated form.

In use case 3 (Energy smart meters) no endpoint validation is performed. Data is agreggated and data produced by these endpoints is always used in correlation with sources with higher confidence levels. For example, if the trust level of the data is medium or low it is not used for taking actions that can have an impact on the operation of the grid.

Source filtering

- Use devices with authentication capabilities to ensure that validation of endpoint sources is possible
- Assign confidence levels on the endpoint sources
- Re-evaluate confidence levels of the endpoints regularly, specially after patches or changes in firmware
- If confidence in endpoint source is low, use it in combination with other higher confidence endpoint sources for taking actions

Access control and authentication: A key challenge to Big Data systems is ensuring access control to authorized users and entities. Due to the several layers and data stored, a role based model is usually adopted to ensure controlled access. With new technologies like Internet of Things (IoT), devices like sensors and appliances also need access to data. Devices such as sensors and point-of-sale terminals do not have the processing power nor the capabilities to offer advanced access controls. So being part of the ecosystem and connected to the network might lead to access mechanisms that can easy be overpassed. One mitigation measure is to use only devices which can provide the needed level of authentication or use third party products that provide this functionality.

In the SIEM use case the roles and authorizations were provided by the Big Data System. In the second use case Energy smart meters authentication and authorization is provided by a third party system, which required additional configuration in the Big Data System.

The Telecom use case the Big Data system needs to provide user based roles and access, as in this case the Telecom is considered a provider. The product offered by the Telecom, required additional development through a third party system In order to gurantee that only authorized people and entities have access to the right data.

Access control and authentication

- Use authentication and authorization to ensure that Big Data queries are executed by authorized users and entities only
- Use components in the Big Data system that follow same security standards to maintain the desired level of security

Monitoring and logging: The use of non-relational data bases may create other security issues due to lack of capabilities in logging or data tagging, as well as classification. When initially creating NoSQL databases, for managing large amounts of data, the idea of security, confidentiality and authentication were not always taken into account. Because the lack of security was not thought about as part of their nature, additional security products are coming into play to offer the extra functionalities. This is also valid for the case of Big Data providers – users of Big Data systems will want to know how a given computation was done and who has access to the results of it.

Another problem that arises from the large volume of transactions being done in the Big Data system is that the logs themselves become too much and too many. One mitigation measure is to keep only certain logs as required by the user. Another is to store logs compressed and use only metadata extracted from them. When needed only the required logs can be decompressed and used.

In use case 1 (Finance) the security of the non-relational databases were provided by the Big Data System. In the two other use cases (Telecom and Energy smart meters) non-relational databases were secured by a third party system, through additional configuration in the Big Data System.

Big Data monitoring and logging

- Enable logging on nodes participating in the Big Data computation
- Enable logging on databases (relational or not) , as well as Big Data applications
- Detect and prevent modification of logs
- Regularly test the restoration of Big Data backups considering the vast amount of data being used in the system

Following our analysis the table represent the challenges associated with their mitigation measures:

CHALLENGES/ MITIGATION MEASURES	ENCRYPTION	SECURITY TESTING AND CODE AUDITING	CERTIFICATION STANDARDS	RISK ASSESSEMENT	SOURCE FILTERING	ACCESS CONTROL AND AUTHENTICATION	MONITORING AND LOGGING
Source validation and Filtering	Yes	Yes		Yes	Yes	Yes	Yes
Secure computation	Yes	Yes	Yes		Yes	Yes	Yes
Access control and authentication	Yes	Yes	Yes			Yes	Yes
Secure Data Management	Yes	Yes				Yes	Yes
Infrastructure security		Yes	Yes		Yes	Yes	Yes
Supply chain security		Yes	Yes	Yes			
Application software security		Yes	Yes			Yes	
Trustworthiness of devices		Yes	Yes			Yes	
Interoperability of applications		Yes	Yes			Yes	
Secure Use of Cloud computing	Yes	Yes	Yes	Yes		Yes	Yes
Distributed Denial of Service Attacks		Yes			Yes		Yes

From our analysis it seems that Security Testing Procedures could be one of the mitigation measure that is key for ensuring proper security controls. Authentication mechanisms, and use of devices and applications with secure standards or certification, are the next key mitigation measures that must be employed.

5. Recommendations

In this section this report provides a list of recommendations addressed to the organizations considering to adopt or that already have adopted Big Data systems based on the analysis of the security challenges and use cases described above.

Recommendation 1: Policy makers should focus on providing guidance for secure use of Big Data systems in the critical sectors.

Due to the distributed nature of this business model and to the interdependencies with other systems, especially in the case Big Data support critical systems, the policy makers and all experts involved in publishing guidelines should take a holistic approach towards Big Data security. When the implementation supports systems vital to the well-being of the society then the security threat analysis and risk assessment should be done in each component separately follow a top-down approach.

Recommendation 2: Big Data providers or vendors should invest in compliance with security standards for their products (devices, services, cloud etc).

Thustworthiness of devices, systems, end point user machines, and cloud services data, is the biggest challenge Big Data is facing from security perspective; compliance with security standards is one of the mitigation measure that can eliminate this threat. In order to apply also to SMEs, certification schemes have become more flexible and cost efficient offering gradual solutions (self attestation, self certification etc).

Recommendation 3: The competent authorities of the critical sectors should encourage vendors to offer security authentication mechanisms and protocols in their products.

More and more devices are becoming part of the cyber-physical world. These devices require access to data in order to take action. Using devices which does not have the capabilities to provide necessary secure authentication mechanisms and protocols, could make the level of security unacceptable. Encouraging vendors and industry to use devices and applications with such capabilities will help make the Big Data system more secure.

Recommendation 4: The standardisation bodies should adapt existing or create new security standards for Big Data.

Currently there are no certifications used for Big Data. Adapting or creating standards will help the industry to grow and provide better service to the users. In order to do that the standartizatoin bodies should create groups of industry, (Big Data provider and users) as well as include regulators in from the sectors affected and agree on common standards and certifications.

Recommendations 5: Industry players and vendors should invest more into enhancing technical security skills of the staff on Big Data through trainings and certifications.

Due to the potential growth in Big Data systems in the coming years, more technical staff with such skills will be required. Investing in traingings and certification of staff to use and create secure Big Data systems should be employed by the industry players, to help the shortage of people in the area.

References

Tankard, C. (2012) 'Big data security', Network Security, 2012(7), pp. 5–8. doi: 10.1016/S1353-4858(12)70063-6.

Okman, L., Gal-Oz, N., Gonen, Y., Gudes, E. and Abramov, J. (2011) Security Issues in NoSQL Databases, 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE, pp. 541–547. doi: 10.1109/TrustCom.2011.70.

Nance, C., Losser, T., Iype, R. and Harmon, G. (2013) 'Nosql vs rdbms-why there is room for both'.

Bloomberg, J. (2013) The big data long tail [Online/Blog] Available from: http://www.devx.com/blog/the-big-data-long-tail.html (Accessed: 26 September 2014).
CEOS (2011) Data life cycle models and concepts [Online]. Available from: http://wgiss.ceos.org/dsig/whitepapers/Data%20Lifecycle%20Models%20and%20Concepts%20v8.docx (Accessed: 26 September 2014).

Demchenko, Y., Ngo, C. & Membrey, P. (2013) Architecture framework and components for the big data ecosystem, draft version 0.2 [Online]. Available from: http://www.uazone.org/demch/worksinprogress/sne-2013-02-techreport-bdaf-draft02.pdf (Accessed: 26 September 2014).

Demchenko, Y., Membrey, P., Grosso, P. & de Laat, C. (2013a) 'Addressing big data issues in scientific data infrastructure', 2013 International Conference on Collaboration Technologies and Systems (CTS), May 20-24, San Diego. Piscataway: IEEE, pp.48-55.

Demchenko, Y. et al. (2013b) 'Intercloud architecture framework for heterogeneous multi-provider cloud based infrastructure services provisioning', International Journal of Next-Generation Computing, 4 (2).

Dumbill, E. (2012) What is big data? An introduction to the big data landscape [Online]. Available from: http://strata.oreilly.com/2012/01/what-is-big-data.html (Accessed: 26 September 2014).

European Union (2010) Riding the wave: how Europe can gain from the rising tide of scientific data [Online]. Available from: http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf (Accessed: 26September 2014).

Gartner (2011) Big data [Online]. Available from: http://www.gartner.com/it-glossary/big-data/ (Accessed: 26 September 2014).

Gray, J. (2010) 'Jim Gray on escience: a transformed scientific method'. In: Hey, T., Tansley, S. & Tolle, K. (eds.) The fourth paradigm: data-intensive scientific discovery. Redmond: Microsoft Research, pp.xvii-xxxi

[Online]. Available from: http://research.microsoft.com/en-us/collaboration/fourthparadigm/ (Accessed: 26 September 2014).

Gualtieri, M. (2013) The Forrester wave: big data predictive analytics solutions, q1 2013. January 3, 2013. [Online] Available from: https://www.forrester.com/The+Forrester+Wave+Big+Data+Predictive+Analytics+Solutions+Q1+2013/fulltext/-/E-RES85601?objectid=RES85601 (Accessed: 26 September 2014).

Liu, F. et al (2011) NIST cloud computing reference architecture [Online]. Available from: http://collaborate.nist.gov/twiki-cloud-computing/pub/CloudComputing/ReferenceArchitectureTaxonomy/NIST_SP_500-292_-_090611.pdf (Accessed: 26 September 2014).
Manyika, J. et al. (2011) Big data: the next frontier for innovation, competition, and productivity [Online]. Available from: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (Accessed: 26 September 2014).

NIST (2013a) NIST big data reference architecture, draft version 1.2. [Online]. Available from: http://bigdatawg.nist.gov/_uploadfiles/M0226_v10_1554566513.docx (Accessed: 26 September 2014).

NIST (2013b) NIST big data working group (NBD-WG) [Online]. Available from: http://bigdatawg.nist.gov/home.php (Accessed: 26 September 2014).

Thanos, C. (2010) Global research data infrastructures: towards a 10-year vision for global research data infrastructures, final roadmap [Online]. Available from: http://www.grdi2020.eu/Repository/FileScaricati/6bdc07fb-b21d-4b90-81d4-d909fdb96b87.pdf (Accessed: 26 September 2014).

Villars, R.L., Olofson, C.W. & Eastwood, M. (2011) Big data: what it is and why you should care. June 2011. [Online]. Available from: http://sites.amd.com/us/Documents/IDC_AMD_Big_Data_Whitepaper.pdf (Accessed: 26 September 2014).

Wladawsky-Berger, I. (2013) Reflections on big data, data science and related subjects [Online/Blog] Available from: http://blog.irvingwb.com/blog/2013/01/reflections-on-big-data-data-science-and-related-subjects.html (Accessed: 26 September 2014).

Shtern, M., Simmons, B., Smit, M. and Litoiu, M. (2013) Toward an Ecosystem for Precision Sharing of Segmented Big Data, 2013 IEEE 6th International Conference on Cloud Computing (CLOUD). IEEE, pp. 335–342. doi: 10.1109/CLOUD.2013.131.

Goel, V. (2014). Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry. [online] Nytimes.com. Available at: http://www.nytimes.com/2014/06/30/technology/facebooktinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html?_r=0 [Accessed 1 Mar.

2015].

Tene, O. and Polonetsky, J. (2012) '11 Northwestern Journal of Technology and Intellectual Property 2012-2013 Big Data for All: Privacy and User Control in the Age of Analytics', Nw J Tech & Intell Prop.

ENISA

European Union Agency for Network
and Information Security
Science and Technology Park of Crete (ITE)
Vassilika Vouton, 700 13, Heraklion, Greece

Athens Office

1 Vass. Sofias & Meg. Alexandrou
Marousi 151 24, Athens, Greece

Catalogue Number TP-02-15-863-EN-N

PO Box 1309, 710 01 Heraklion, Greece
Tel: +30 28 14 40 9710
info@enisa.europa.eu
www.enisa.europa.eu

ISBN: 978-92-9204-142-7
DOI: 10.2824/13094