Big Data Security And Privacy

Muhammad Danish
School of Software, Xinjiang University, Urumqi 830008, China

Abstract

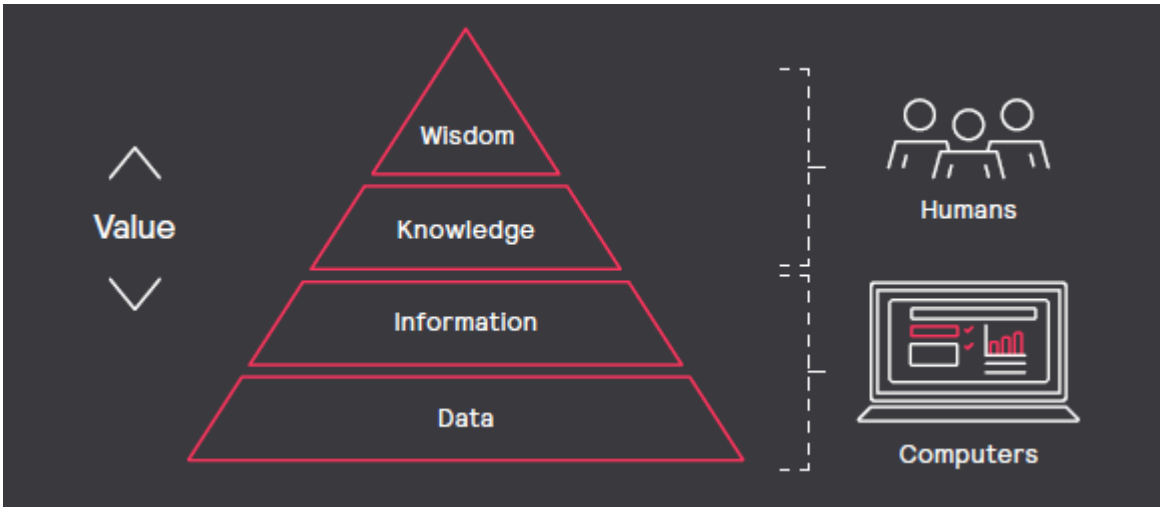
Big knowledge (DATA) has emerged as a necessity within the gift world. Most of the persons are connected to at least one another through completely different modes of communications. Folks share data in numerous forms. The knowledge that connects folks is growing staggeringly in giant volume that is making security and privacy considerations. As massive knowledge technologies are rising at no time pace, it's conjointly making the house for security and privacy problems. Till these problems don't seem to be self-addressed properly, it's going to produce obstacles to the fulfillment of expected growth and opportunities and long-run success of massive knowledge. During this paper, we have a tendency to review the assorted domains of massive knowledge like tending, social media, net of things (IOT) and social networking for security and privacy connected problems. Thus protective its security is extremely vital and changing into a prime priority for several organizations. Sadly there's no single formula which will guarantee 100% of data security. Thus there's a desire for a group of benchmarks or standards to make sure the most effective security practices area unit adopted and an adequate level of security is earned. During this paper, authors introduce varied info security standards shortly and so give a comparative study for major info security standards.

Keywords — ***Big*** data, security, privacy, data knowledge, data sharing**.**

I. INTRODUCTION

Smart connected infrastructures, whether or not in cities, workplaces, retail environments, or homes, generate giant amounts of knowledge. Municipalities and privately-owned businesses will gain the competitive advantage by reducing service delivery prices and streamlining operations; anticipating and satisfying the requirements of voters, employees, and customers; and making new revenue streams through new data-driven merchandise and services. Organizations will mix knowledge gathered through connected applications and infrastructures with relevant external knowledge and domain models to realize a deeper understanding of people's behavior and interactions. (The Internet of Things, 2019)

In the most general terms, knowledge could be an assortment of numbers or characters. Knowledge may be measured, collected, analyzed and given victimization numerous formats like tables, carets, graphs, and images. Conceptually, knowledge refers to the very fact that some existing info or information is depicted or coded in some type appropriate for higher usage or process. Data, info, knowledge, and knowledge area unit closely connected ideas, however, everyone is distinct. Knowledge alone has a very little price because it should be processed and contextualized to yield unjust insights. A widely used model to represent relationships among knowledge, information, knowledge, and knowledge is thought because of the DIKW pyramid. (Manyika., James., 2015).



This model represents a hierarchy, and implies a series of transformations for ascending the hierarchy. Information is that the basic constituent component of data. To become data, information should be processed. Processed and understood data should be analyzed to yield information. Principles should be applied to information to lead to knowledge. As a result of knowledge involves victimization information and skill for the bigger sensible or a high-level goal, it's deeper and a lot of unambiguously human than information. It needs a way of fine and dangerous, right and wrong, moral and unethical. It involves Associate in Nursing understanding of individuals, objects, events, and things, and also the disposition in addition because the ability to use perception, judgment, action to keep with an understanding of the best course of action. As such, knowledge is in and of itself subjective. As the DIKW pyramid shows, information derives that means Associate in Nursing price from process with a purpose with an finish goal and a business objective in mind. For businesses, this needs Associate in Nursing structure specialize in information management processes. (Schneier, Bruce,2016).

II. DATA CATEGORIES: STRUCTURE, SIZE, SPEED, AND SOURCE

It is necessary to understand that there square measure totally completely different classes of knowledge that have different management needs. the most classes square measure structure, size, speed, and source. every of those should be accounted for in an efficient knowledge management method. Technics Publications, 2017.

The data structure is defined into three categories are below:

A. Structured

Knowledge refers to knowledge with a high-level organization. Structured knowledge incorporates a pre-defined knowledge model or a schema a model of the kinds of knowledge which will be recorded and the way they'll behold on, processed, and accessed. The schema includes process what knowledge is going to behold on and the way it'll behold on, together with knowledge kind (numeric, currency, alphabetic, name, date, address) and any knowledge input restrictions (number of characters, specific terms, numeric ranges). Technics Publications, 2017

B. Semi-Structured

Knowledge could be a style of structured knowledge that lacks a strict knowledge model structure. With semi-structured knowledge, tags or alternative forms of markers square measure wont to establish sure among the info, however, the info itself doesn't have a rigid structure. As an example, emails have the sender, recipient, date, time and alternative fastened fields added to the unstructured knowledge of the e-mail message content and any attachments. Photos or alternative graphics will be labeled with keywords like the creator, date, location, and keywords, creating it doable to arrange and find graphics. File systems and varied file formats square measure typically wont to manage semi-structured knowledge. Technics Publications, 2017

C. Unstructured

Knowledge refers to knowledge that either doesn't have a pre-defined knowledge model or isn't organized in an exceedingly pre-defined manner. Unstructured knowledge is usually text-heavy, however, might contain alternative forms of knowledge like dates and numbers. This leads to irregularities and ambiguities that build it tough to know exploitation ancient programs as compared to structured knowledge. Techniques like data processing, tongue process (NLP), and text analytics offer totally different strategies to search out patterns in, or otherwise interpret, unstructured data, samples of unstructured knowledge square measure social media, web page, and center logs. Technics Publications, 2017

III. SIZE: SMALL DATA AND BIG DATA

Another way to reason knowledge is by its size. The knowledge that may slot in the memory of a laptop/ computer and may be managed by ancient processing applications is termed little, whereas knowledge that's therefore massive or complicated that it cannot be restrained mistreatment ancient processing applications are termed massive knowledge. Ancient relational knowledge base management systems and desktop statistics and visualization packages usually have an issue handling massive data. Process massive knowledge could need massively parallel package running on tens, hundreds, or maybe thousands of servers. What counts as “big data” varieties betting on the capabilities of the users and their tools and increasing capabilities, creating process massive knowledge a moving target. (Wired,2017)

IV. SPEED: DATA AT REST, DATA IN MOTION, SLOW DATA, AND FAST DATA

Another way to reason knowledge is by its dynamic characteristic. The knowledge that's static in nature, i.e. keep in persistent storage (disk, type) in any digital type (e.g. database, knowledge warehouse, computer program, files) is usually known as knowledge at rest. Knowledge inmotion is

that the term used for knowledge because it is in transit. Knowledge in motion involves the process of knowledge on the fly while not storing it. A typical characteristic of knowledge in motion is the rate. The rate is that the rate of flow at that the info is formed, stored, analyzed, and visualized. Quick knowledge rate means that knowledge is being processed in a very short quantity of your time. Within the quick huge knowledge era, knowledge is formed and passed on in real time or close to real time. Increasing knowledge flow rates produce new challenges to change real- or close to a period of time knowledge usage. (Banafa, Ahmed, 2017)

Traditionally, this idea has been delineated as streaming knowledge. The second characteristic for knowledge in motion is variability that refers to any amendment in knowledge over time, together with the flow, the format, or the composition. Provided that several knowledge processes generate a surge within the quantity of knowledge incoming in a very given quantity of your time, new techniques square measure required to expeditiously handle this knowledge. (Banafa, Ahmed, 2017)

V. SOURCE: INTERNAL DATA AND EXTERNAL DATA

From the information supply or origin perspective, information may be categorized as internal or external. Information that comes from internal systems (for example, company systems, IT applications) is termed internal, whereas information that comes from third-party services is termed external, samples of external information area unit social media, whether or traffic information from third parity net services. (Banafa, Ahmed, 2017)

Traditional enterprise info management systems take care of structured or at the most semi-structured information, tiny and slow information, information at rest and internal information. IoT data, that area unit usually characterized by unstructured, huge information streams and external information, demands new architectures for managing such information. (Banafa, Ahmed, 2017)

Data management is vital to making sure security and privacy within the growing IoT and sensible systems area, which has connected lighting systems for cities, workplaces, retail environments, and homes. Information management includes all the disciplines required to firmly manage information as a valuable plus. Palm-information management needs an information management design that permits a unified period of time reading of all data assets, that area unit generally derived from disparate sources. Ideally, information management design manages associate degree organization's information assets throughout its lifecycle, from the purpose of acquisition to consumption in end-user applications and on to deletion. Central governance and a corporation-wide information management strategy area unit required to align data-related activities and maximize the worth of information assets. (Banafa, Ahmed, 2017).

Bellow steps are to maintain successful data management:

Manage the data lifecycle

Provide access to all acquired data

Ensure data quality

Provide a unified view of all data assets to enable integration across system verticals and enterprise silos

Include processing, transformation, and enrichment of data.

Specify data governance policies and procedures to ensure the highest possible level of data security and privacy

VI. DATA MANAGEMENT IN CONCEPT

Data management style manages data assets from acquisition to consumption. as a result of the subsequent figure illustrates, abstract data management style at the structure level for a city or associate enterprise are usually sophisticated, comprising multiple functions and capabilities. data governance, they operate primarily answerable for mitigating risk, implementing compliance, and making sure security, can be one vertical operate that touches all layers of the planning. (Big Data, 2013)

Data acquisition: The ability to amass information of varied classes (structure, size, speed) from a range of sources (IoT, enterprise, social media).

Data cloth: The routing of knowledge from acquisition to storage, end-use application, or both. This operation can even embrace an information enrichment pipeline to reinforce data quality.

Data storage: The power to store multiple information varieties for various desires, as well as a storage sink or country for initial temporary storage, a knowledge lake or reservoir for economically storing large scale, gently structured information, an information} warehouse for primary structured data storage, associated an enterprise information warehouse for important enterprise information storage.

Data management: The ability to manage information across multiple information stores, perform ancient and untraditional extract rework and cargo (ETL) operations, and implement governance, as well as information provisioning, lifecycle management, and information quality functions.

Data access: The access layer that gives information routing to and from applications, and information virtualization if required.

Data analysis and process: The applying layer for consumption of information (big data sandbox discovery, business analytics, information modeling, and transformation).

Governance, risk, compliance, and security: Distributed practicality for implementation of all information management and security processes. This abstract information management design isn't implementation-specific, and it doesn't address problems associated with personal cloud, public cloud, on-premise, or hybrid implementations.

Data acquisition: the power to amass information of varied classes (structure, size, speed) from a range of sources (IoT, enterprise, social media).

Data cloth: The routing of knowledge from acquisition to storage, end-use application, or both. This operation can even embrace an information enrichment pipeline to reinforce data quality.

Data storage: The power to store multiple information varieties for various desires, as well as a storage sink or country for initial temporary storage, a knowledge lake or reservoir for economically storing large scale, gently structured information, an information warehouse for primary structured data storage, associated an enterprise information warehouse for important enterprise information storage. (Fayyad, 2019)

VII. PRIVACY

Privacy is expounded to security and compliance however has its own definition, risks, and mitigation methods. Privacy typically applies to a consumer's or user's right to safeguarded personal data from use by others. Probably vulnerable information includes, however, isn't restricted to, information shared on social media and any reasonably demographic or personal information. in an exceedingly sensible system, vulnerable information might embrace information a couple of worker's habits (in and out time, regular payment and alternative unit of time data); a citizen's preferences, engagement in crowdsourcing applications, or movements around a city; or a shopper's shopping for habits, checking account, and credit data. In general, privacy is Associate in nursing individual's right to stay this and alternative forms of personal information to herself. (Villars, 2011)

One main goal of security is the protection of Associate in Nursing enterprise, organization, or agency which can or might not store and manage vulnerable personal information. Whereas privacy and security, objectives will typically coincide, security policies and procedures might not address all privacy issues. for instance, business or municipality could secure the private information it stores from cyber-attacks, however, staff or officers within the network could also be ready to review this information. A really common situation is one during which an internet distributor has top-of-the-line system security measures in situ, however, it freely sells personal information to understand secondary revenue streams. (Patil and Seshadri, 2014).

A. Private Information, Confidential Information, Open Data

Privacy and confidentiality area unit closely connected. Privacy is sometimes understood to sit down with Associate in Nursing individual's information (private to the person), whereas confidentiality is sometimes understood to sit down with a company's information (private to the organization). Terribly usually, we will point out the connection of collected information to privacy and confidentiality: Privacy-related information in person acknowledgeable and in a person sensitive information that contains privacy-invasive data restricted by law and moral conditions. (Middleton, Kjeldsen and Tully, 2013)

a) Confidential information and trade secrets:

Business-related information like methods, blueprints, formulas/ recipes, and operational processes. Could or might not be totally or partly restricted looking on a company's policies and agreements with partners and alternative third parties. (Groves, Kayyali, Knott, and Kuiken, 2013).

b) Open data

Information with no privacy or confidentiality problems. Such information could also be collected and shared while not restrictions.

B. Legal and moral risks:

Legal and moral risks accompany the assembling, monitoring, processing, and storing of knowledge derived from sensible systems. In the least times, there ought to be Associate in a Nursing unjust assessment of the potential advantages of collected information versus the potential harms that the gathering, storage, and use of such information might wearer the privacy of people and teams, still as on the moral norms and standards of society. These issues hold for all affected constituencies, whether or not voters in an exceeding municipality, staff in an exceedingly geographic point, customers in an exceedingly retail atmosphere, or residents reception.

Privacy issues in sensible cities area unit particularly acute, providing misuse of collected information could result in abuses of power and

discrimination that might undermine the fundamental principle of human equality and result in unwelcome dependency on those that management the info.

C. Challenges to information security and privacy

The increase of devices connected to the net and connected to each totally different, the degree of data collected, stored, and processed is increasing every day that together brings new challenges in terms of the info security. In fact, the presently used security mechanisms like firewalls that cannot be utilized within the huge info infrastructure as a result of the protection mechanisms got to be stretched off the perimeter of the organization's network to satisfy the user/data quality desires and conjointly the policies of BYOD (Bring Your Own Device). Considering these new things, the pertinent question is what security and privacy policies and technologies area unit further up to satisfy this prime huge info privacy and security demands. These challenges may even be organized into four huge info aspects like infrastructure security, info privacy (e.g. processing that preserves privacy/granular access), info management (e.g. secure info starting and storage) and, integrity and reactive security (e.g. real-time looking of anomalies and attacks). Considering huge info there is an assortment of risk areas that need to be thought of. These embrace the info lifecycle (provenance, possession, and classification of data), the data creation and assortment technique, and conjointly the shortage of security procedures. Ultimately, the huge info security objectives don't seem to be any completely totally different from the opposite info kinds to preserve its confidentiality, integrity, and convenience. (Ranjan, Wang, Khan, Zomaya, 2015)

Being huge info such an important and complicated topic, it's nearly natural that security and privacy challenges will aeries. huge info has specific characteristics that have a bearing on data security: choice, volume, velocity, value, variability, and honesty. These challenges have a direct impact on the look of security solutions that area unit required to tackle those characteristics and wishes. Currently, such out of the box security resolution does not exist. (Sundmaeker, 2010)

Cloud Secure Alliance (CSA), a non-profit organization with a mission to push the employment of best practices for providing security assurance within Cloud Computing, has created enormous info operating parity that has focused on the foremost necessary challenges to implement secure massive info services. CSA has classified the wholly totally completely different security and privacy challenges into four different aspects of the large system. Each next security challenges as per CSA are below:

- ***Infrastructure Security***

Secure Distributed process of knowledge
Security Best Actions for Non-Relational Data-Bases

- ***Data Privacy***

Information Analysis through data processing protective information Privacy
Scientific discipline Solutions for Information Security
Granulate Access management

- ***Data Management and Integrity***

Secure information Storage and dealing Logs
Granulate Audits
Information beginning

- ***Reactive Security***

End-to-End Filtering & Validation
The direction the protection Level in a time period
The increase inside the variability of connected devices (cares, lighting systems, refrigerators, telephones, glasses, management systems, health looking at devices, TVs, home security systems, home automation systems, and plenty of more) has LED to manufacturers to push to the market, in Associate in Nursing extremely short quantity of it slow, Associate in Nursing outsize set of devices, cloud systems and mobile applications to use this opportunity. Whereas it presents tremendous benefits and opportunities for end-users it's conjointly chargeable for security challenges. (Checkmarx, 2017)

There are top 10 known the subsequent security problems are below:

1) Insecure web Interface: which could allow the associate aggressor to use associate administration web interface and acquire unauthorized access to control the IoT device.

2) deficient Authentication/Authorization: can allow the associate aggressor to use a foul parole policy, break weak passwords and access to privileged modes on the IoT device.

3) Insecure Network Services: The cause could associate aggressor exploiting spare or weak services running on the device, or use those services as a jumping purpose to attack totally different devices on the IoT network.

4) Lack of Transport Encryption: allowing the associate aggressor to pay attention info in transit between IoT devices and support systems.

5) Privacy Concerns: It concern from the particular truth the foremost IoT devices and support systems

collect personal info from users and fail to defend that info.

6) Insecure Cloud Interface: whereas not correct security controls associate aggressor can use multiple attack vectors (insufficient authentication, lack of transport secret writing, account enumeration) to access info or controls via the cloud electronic computer.

7) Insecure Mobile Interface: whereas not correct security controls associate aggressor can use multiple attack vectors (insufficient authentication, lack of transport secret writing, account enumeration) to access info or controls via the mobile interface.

8) Deficient Security Configurability: as a result of the dearth or poor configuration mechanisms associate aggressor can access info or controls on the device.

9) Insecure Software/Firmware: attackers can profit of unencrypted and unauthenticated connections to hijack IoT devices updates, and perform a malicious update which is able to compromise the device, a network of devices and additionally the data they hold.

10) Poor Physical Security: if the IoT device is physically accessible than associate aggressor can use USB ports, American state cared or totally different storage means to access the device OS and doubtless any info hold on the device.

It is cleared that huge info gift gripping chance for users and businesses, however, these opportunities area unit countered by vast challenges in terms of privacy and security. (Checkmarx, 2017)

VIII. RECOMMENDATION FOR MASSIVE INFORMATION SECURITY AND PRIVACY

There is no single charming resolution to resolve the proverbial huge data security and privacy challenges and ancient security solutions, that are primarily dedicated to defending small amounts of static data, are not capable the novel requisites obligatory by huge data services. There's the need to know but the gathering of giant amounts of sophisticated structured and unstructured data are usually protected. Non-authorized access to that data to create new relations, combine altogether totally different data sources and build it accessible to malicious users could also be a heavy risk for big data. the elemental and extra common resolution for this includes encrypting everything to form data secure regardless of where the knowledge resides (data center, computer, mobile device, or any other). (Mohamed, Ahmad, 2017)

Due to its characteristics, huge data comes have to be compelled to take associate holistic vision at security. Huge data comes to have to be compelled to take into thought the identification of the assorted data sources, the origin, and creators of information, however as UN agency is allowed to access the knowledge. It's put together necessary to conduct associate degree correct. As a recommendation, altogether totally different security mechanisms have to be compelled to be nearer to the {data} sources and data itself, therefore on providing security right at the origin of information, and mechanisms of management and clean on archiving, data leak clean and access management have to be compelled to work on.

The new huge data security solutions have to be compelled to extend the security perimeter from the enterprise to the overall public cloud. Throughout this technique, a trusting data starting mechanism has to be compelled to be put together created across domains. To boot, similar mechanisms to those utilized in are usually used to mitigate distributed denial-of-service (DoS) attacks launched against huge data infrastructures. Also, huge data security and privacy are vital to verify data attribute throughout the whole data lifecycle from data assortment to usage.

While creating an effort to want the foremost of giant data, in terms of security and privacy, it becomes necessary that mechanisms that address legal desires regarding data handling, have to be compelled to be met. Secure secret writing technology ought to use to defend all the direction. Personally identifiable info (PII), Protected Health data (PHI) and property (IP) and careful science material (keys) access management policies, have to be compelled to be placed in place, to verify the right protection and unlocking information this can be often notably necessary for data hold on. Therefore on win success, these mechanisms have to be compelled to be clear to the end-user and have a low impact of the performance and quality of information code and hardware-based encryptions mechanisms are to be thought about.

Important security and privacy challenge for big data are expounded with the storage and method of encrypted data. Running queries against associate encrypted data could also be a basic security demand for secure huge data however it is a troublesome one. This raises queries like a) is that the knowledge encrypted with one or multiple keys; b) can the knowledge needs to be rewritten before running the query; c) do the queries have to be compelled to be put together encrypted; d) UN agency as a result of

the permissions to decode the database; and much of extra. SupportedCrypt DB, Google has developed the Encrypted huge question shopper which will allow encrypted huge queries against their huge question service that permits super, SQL-like queries against append-only tables, practice the method power of Google's infrastructure. (Patil and Seshadri, 2014)

Apart from extra specific security recommendations, it's put together necessary to ponder the protection of the IT infrastructure itself. one all told the common security practices is to position security controls at the sting of the networks, however, if associate aggressor violates this security perimeter it's going to have access to all or any or any the knowledge within it. Therefore, a replacement approach is vital to maneuver those security controls with regards to the knowledge (or add additional ones). Monitoring, analyzing and learning from data usage and access is to boot an important aspect to unceasingly improve the security of the knowledge holding infrastructure and leverage the already existing security solutions.

IX. REFERENCES

[1]. Big Data Security And Privacy Sponsored by the National Science Foundation, The University of Texas at Dallas, September 16-17, 2014

[2]. Ira S. Rubinstein, 2013, Big Data The End of Privacy or a New Beginning?

[3]. U M Fayyad, 'From Data Mining to Knowledge Discovery, An Overview,' in U M Fayyad, Advances in Knowledge Discovery and Data Mining 6 (Menlo Park: AAAI, 1996), cited in Tal Z Zarsky, 'Mine Your Own Business! Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion' (2003) 5 Yale Journal of Law and Technology.

[4]. McKinsey Global Institute, 2011 'Big Data: The Next Frontier for Innovation, Competition, and Productivity'.

[5]. John Markoff, 2012 'How Many Computers to Identify a Cat? 16,000' NY Times B1,

[6]. ShashiRekha .H., Dr. Chetana Prakash, Kavitha G., 2014 "Understanding Trust and Privacy of Big Data in Social Networks: A Brief Review "

[7]. Richard L Villars, Mathew East Wood, and Carl W. Olofson, 2011 "Big Data what it is and why you should care".

[8]. Won Kim, Ok-ran Jeong, Sang-won lee. 2009 "On social websites".

[9]. Harsh KupwadePatil and Ravi Seshadri, 2014 'Big data security and privacy issues in healthcare',

[10]. P. Institute, 2012 "Third Annual Benchmark Study on Patient Privacy and Data Security," Ponemon Institute LLC.

[11]. P. Middleton, P. Kjeldsen and J. Tully, 2013 "Forecast: The Internet of Things, Worldwide," Gartner.

[12]. M. Hanson, H. Powell, A. Barth, K. Ringgenberg, B. Calhoun, J. Aylor and J. Lach, 2009 "Body Area Sensor Networks: Challenges and Opportunities," Computer, pp. 58-65.

[13]. P. Groves, B. Kayyali, D. Knott and S. V. Kuiken, 2013 "The 'big data' revolution in healthcare," McKinsey & Company.

[14]. Boyd and K. Crawford. 2011 Six Provocations for Big Data. SSRN eLibrary.

[15]. Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, Michele Zorzi, 2014 "Internet of Things for Smart Cities" IEEE Internet Of Things Journal, Vol. 1, No. 1,

[16]. C. Perera et al., "Context-Aware Computing for the Internet of Things: A Survey," IEEE Comm. Surveys & Tutorials, vol. 16, no. 1, 2013, pp. 414–454

[17]. C.Perera R. Ranjan, Lizhe Wang; S.U. Khan, A.Y. Zomaya, 2015 "Big Data Privacy in the Internet of Things Era", IT Pro May/June.

[18]. H. Sundmaeker et al., "Vision and Challenges for Realizing the Internet of Things," Cluster of European Research Projects on the Internet of Things, 2010, www.internet-of-things -research.eu/ pdf/IOT_Clusterbook_March_ 2010.pdf.

[19]. Delphine Christin, Pablo Sanchez Lopez, Andreas Reinhardt, Matthias Hollick and Michael Kauer" 2012 Share with Strangers: Privacy Bubble as user-centered privacy control for mobile content sharing applications "Elsevier.

[20]. Norshidah Mohamed, Eli Hawa Ahmad "Information Privacy concerns, antecedents and privacy measure use in social networking sites: Evidence from Malaysia

[21]. "The Internet of Things." Extracted on 18th March 2019, IT Glossary, Gartner: Retrieved from http://www.gartner.com/it-glossary/internet-of-things/Manyika, James, et al. "Unlocking the potential of the Internet of Things."McKinsey Global Institute, June 2015: http://www.mckinsey.com/businessfunctions/digital-mckinsey/our-insights/the-internet-of-things-thevalue-of-digitizing-the-physical-world

[22]. "There will be 24 billion IoT devices installed on Earth by 2020." Business Insider, Extracted on 9 June 2016: http://www.businessinsider.com/there-will-be-34-billion-iot-devices-installed-on-earth-by-2020-2016-5

[23]. Schneier, Bruce. "Data Is a Toxic Asset." Schneier on Security, 4 March 2016: Retrieved from http://www.schneier.com/blog/archives/2016/03/data_is_a_toxic.html

[24]. Bauer, Harald, Burkacky, Ondrej, and Knochenhauer, Christian. "Security in the Internet of Things." McKinsey & Company, May 2017: http://www.mckinsey.com/industries/semiconductors/ourinsights/security-in-theinternet- of-things

[25]. Nuttall, Nathan, Goodness, Eric, Hung, Mark, and Geschickter, Chet. "Survey Analysis: 2016 Internet of Things Backbone Survey." Gartner, 5 January 2017: http://www.gartner.com/doc/3563218/survey-analysis--internet-things

[26]. DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition). Technics Publications, 2017: p. 17.

[27]. Huijbregts, Rick. "Re-imagining business value in a digital world." Cisco, May 2016.

[28]. "The biggest security threats coming in 2017." Wired, 2 January 2017: http://www.wired.com/2017/01/biggest-security-threats-coming-2017/

[29]. Doctorow, Cory. "Winter Denial of Service attack knocks out heating in Finnish homes." BoingBoing, 8 November 2016: http://boingboing. net/2016/11/08/winter-denial-of-service-attac.html

[30]. Banafa, Ahmed. "Three Major Challenges Facing IoT." SemiWiki.com, 25 May 2017: http://www.semiwiki.com/forum/content/6796-three-majorchallenges- facing-iot.html

[31]. Definitions based in part on the Wikipedia entry on data, at HTTP:// en.wikipedia.org/wiki/Data

[32]. "Data Lifecycle Overview." USGS website: http://www2.usgs.gov/datamanagement/why-dm/lifecycleoverview.php

[33]. Rubens, Arden. "A Closer Look: OWASP Top 10 Application Security Risks." CheckMarx, 22 May 2017: http://www.checkmarx.com/2017/05/22/ closer-look-owasp-top-10-application-security-risks/