Big Data Utilization in Security Sector: A Case Study of the United States of America

By

Ehsan Shayegan

Submitted to

CENTRAL EUROPEAN UNIVERSITY (CEU)

School of Public Policy (SPP)

Supervised by

Dr. Evelyne Hübscher

In Partial Fulfillment of the Requirement for Master of Arts in Public Policy

One Year Program

June 2020

Disclaimer

I Ehsan Shayegan hereby declare that this thesis entitled “Big Data Utilization in Security Sector: A Case Study of the United States of America” is composed solely by me. It is submitted to the Central European University (CEU) in partial fulfillment of the requirement for one-year master program at School of Public Policy (SPP).

The information put forth are based on my understanding from the credible and different sources. It is not previously included partially or fully in any thesis, submitted to CEU or any other institution. This thesis is completed in absolute adherence to the CEU’s enacted research ethic guidelines and policies. Except where the used references are acknowledged, the work presented is entirely my own.

Signature ……

June 2020

Acknowledgement

I want to thank first and foremost my thesis advisor, Dr. Evelyne Huebscher for her insightful, timely and instructive guidance throughout thesis writing phase. Undoubtedly, this work would not have been possible without her constructive and rightful instructions.

I am very grateful to Central European University (CEU), especially to the School of Public Policy (SPP) for granting me this amazing opportunity. Also, to my professors whom I owe greatly because of their professional, honest and tireless efforts in educating me.

Finally, I must express my profound gratitude to my wife Dr. Setara Sahar and to my lovely daughter Adrina Shayegan whom provided me the whole heartedly support, encouragement and inspiration. More importantly standing by me during the tough times; taking care of me with the minimum health facility during my self-isolation as Covid-19 patient in Afghanistan. Undeniably, their generous love made my one-year academic journey possible.

Table of Content

List of Abbreviations

AI	Artificial Intelligence
AIM	Augmenting Intelligence Using Machines
ALPR	Automatic License Plate Reader
BD SSG	Big Data Senior Steering Group
CEU	Central European University
CIA	Central Intelligence Agency
DOD	Department of Defence
DHS	Department of Homeland Security
DNI	Director of National Intelligence
FBI	Federal Bureau of Investigation
HSI	Homeland Security Investigation
IARPA	Intelligence Advanced Research Projects Activity
IC	Intelligence Community
JTTF	Joined Terrorism Task Force
JAIC)	Joint Artificial Intelligence Centre
LAPD	Los Angeles Police Department
LASER	Los Angeles' Strategic Extraction and Restoration program
NSF	National Science Foundation
NSA	National Security Agency
NITRD	Networking and Information Technology Research and Development
NPM	New Public Management
OBP	Office of Border Patrol
OSTP	Office of Science and Technology Policy
PCAST	President's Council of Advisors on Science and Technology
R&D	Research and Development
SPP	School of Public Policy
SCS	Social Credit System
USA	United States of America

Abstract

My thesis goal is to assess the use of big data in U.S security sector. It tries to draw recommendation, on the basis of lessons from U.S big data initiatives in security agencies. This is carried out by relying on secondary data; gathered from different academic journals, books, official reports, strategic plans and online archives of the U.S government.

Based on the results from this study, U.S is one of the successful examples of big data utilization in public sector. Specifically, the deployment of big data in U.S security agencies, has enormously transcended the effectiveness of security measures; enabling them in better prediction, identification, and mitigation of crimes and threats. Also, this study finds that big data utilization has become a prevalent practice for informing the decisions in U.S security agencies.

However, the success in U.S security agencies is understandable in a bigger enabling environment. In addition to the collaboration among sectors, the political and economic factors have greatly contributed to this mission. Despite the optimism in effectiveness of big data utilization in security area, this thesis emphasizes on the importance of data privacy and the necessity of keeping the proper balance between data acquisition and protection of human rights values such as freedom, civil liberties and personal privacy.

CHAPTER I: INTRODUCTION

Definition and Evolution of Big Data

The chronology of the term big data dates back to 1977 used by NASA scientists Michael Cox and David Ellsworth. Later in 1998 it was used by John Masey, an SGI data scientist in a paper titled "Big data and the Next Wave of Infrastress". John's paper tries to elaborate the volume of data, storage and the necessity for analysis tools. He has contributed substantially in defining the term big data which is similar to the current notion of big data (Chan et al. 2018, 3).

According to Chan et al. (2018), in today's debate around big data, there is a huge literature and scholars have defined it from different perspectives. The mutuality among them is referring to big data as a vast amount of data, generated and captured in a variety of formats from a number of disparate sources. This high volume of data exists in structured, semi- structured, and unstructured forms that has outpaced the current ability of standard data tools and methods of analysis (6).

Chan et al. (2018) states that there is no universally consensus on a rigorous definition of big data within the industry; however, the term is generally defined by some major characteristics; such as volume *(hugeness of data)*, velocity *(speed of data processing)*, variety *(disparity of the data sources)* (6), veracity *(disparity and noise in data)* and smart analytics *(combination of analytical tools with higher computational power)*. More importantly the way big data is used for different purposes (3).

Historically the advancement of big data during the last two decades has become unprecedented. Currently, it plays a revolutionary role in overall management of both public and private sector. Though, its emergence is a relatively new phenomenon but the tradition of storing big data dates back to the early 1950s. It has experienced relatively slower pace until mid-1990s, mainly due to the high cost of computer and data storage and data networks (Lee 2017).

Lee (2007) further stated that the big data has been accelerated vastly when in 1994 web mining techniques, web usage mining, web structure mining and web content mining were developing. It was absolutely revolutionized between the years 2005-2014 with the prevalence of social media content mining.

Big data entered in a new phase by the development of data streaming analytics, as the analysis of audio and video has become conceivable phenomenon. By its vast expansion, it has paved its way to several sectors such as agribusiness, financial regulations and medicine. (Lee 2017, 298).

Approximately 90% of the current global data has been created during the past two years, with 2.5 quintillion bytes of data added each day. Interestingly almost 90% of the produced data is unstructured which is usable only by utilization of big data technology (Kim et al. 2014). For better governance, public agencies see the data outburst as a great opportunity. Thus, in recent years the developed nations have invested largely to harness the big data power for their own well-being purposes (Strang et al. 2017).

However, in this study, the term big data is defined as a huge amount of structured, semi-structured and unstructured data that mainly includes documents, natural language, web data, social media data, multi-media data (image, audio, video) and mobile data *(sensor, geographical location and applications).* It further tends to include the use of big data technology, tools and analytics. To be precise, when discussing big data utilization, this study relies only and merely upon the lawful and legitimately accessible data.

Big Data Application

Private sector has pioneered this area but governments are currently embarking to invest on utilization of big data for evidence-based decision making. Thus, all the leading countries in big data area, are finding it as a strategic tool in fulfilling their mission in making the public service easily and equally available, also to effectively involve citizens in public affairs (Kim et al. 2014).

On the application of big data in public sector, a study conducted in 2017, found that from an overall 17% of the entire publication in big data with management approach, 7% of them focuses on organization and public sector management. This study had analyzed the entire publications on big data from 2005 to the year 2017 (Sheng, Amankwah-Amoah and Wang 2017).

Though, it represents a relatively small percentage but by the advancement of technology, big data industry has flourished far beyond our imagination. It has proliferated and prevailed almost every sector including the areas of crime investigation, comprehension, and prediction. Similarly, it is widely practiced in the area of law enforcement and intelligence investigations (Chan et al. 2018, 2).

Kim et al. (2014) studied big data in the area of security. Their study indicated that in some countries big data is prevailing the strategies and tools of fighting terrorism and violent crimes. United States of America (USA) is one of the leading countries in the deployment of big data in its public and security sector and by investing largely in this arena; big data has become one of its national priorities. Apart from other institutions, only the USA's National Science Foundation (NSF) has a dedicated program with $100 million budget in big data category (Strang et al. 2017).

However, despite the prevalence of big data utilization in today's world, the use of big data in security sector is lacking the sufficient and systematic literature. Scholars have emphasized

more on the issues of privacy and civil liberties. Other aspects are yet remaining unexplored. This deficit in the literature is simultaneous with severe implications for the policy makers and researchers, as the absence of a systematic study has posed huge limitations to researchers in understanding the implication of big data advancement in the area of security ((Puyvelde et al. 2017).

This applies particularly to USA as one of the pioneers with several years of tradition in utilizing big data in security sector. Although, a systematic literature in this area could help the policy makers in other countries in understanding the big data promises; which by turn could offer better understanding of the security situation and informing similar policy formulations.

Focus of this Study

This study tends to assess the effectiveness of big data in U.S security sector with emphasis in its potential in identifying, mitigating and predicting the crimes and possible threats. It also looks for evidences that proves big data effectiveness in informing decisions in security agencies. To do so, it assesses the actual scenarios based on what the U.S security strategic plans and policies have envisioned as big data promises. For better understanding of this fact, it tends to posit big data utilization in a bigger environment; examining some major factors that have greatly contributed to advancement of big data utilization. Subsequently, on the basis of lessons from U.S security context, this study tends to conclude by providing a recommendation for countries where such interventions could be transferred as a constructive policy.

To do so, this research relies substantially on secondary data; existing in the forms of journal articles, books, official reports and online archive materials. This thesis encompasses six sections. It begins by providing some generic and conceptual information about big data. In the second chapter, it describes the methodological details. Subsequently in the third chapter, it provides a detailed account of similar literature and the research framework that this study tends to follow. It will be followed by results and analysis chapters, respectively. And finally, it finishes by listing the bibliographical details of all references used in this study.

CHAPTER II: METHODOLOGY

Research Approach and Rational

To undertake this study, I have adapted case study approach. The reason I prioritized this approach is its appropriateness to achieving the goal of this study. It gives me more freedom to reflect on various aspect and analyze security oriented big data utilization in a contextualized and bigger environment.

Sources and Process

This study majorly relies on secondary data; gathered through review of the published literature including books, academic articles, research reports, state reports, policies, strategies and the online White House archive materials. To access them, I used the available online platforms such as online libraries, online journals, credible research institute websites, U.S federal agencies such as Department of Homeland Security (DHS), National Science Foundation (NSF), Federal Bureau of Investigation (FBI), Central Intelligence Agency (CIA) and the White House. Data retrieved from the aforementioned sources are both qualitative and quantitative in nature, however this study is leaning towards the qualitative information.

The sources used for this study are categorized under three different divisions. Firstly, a large chunk of literature is academic in nature. They are used mainly to enlighten the conceptual part, including the definitions, conceptual approach and the relevant produced works by other scholars.

The second categories are semi conceptual and operational documents. This category comprises a wide variety of policies, strategies, guidelines and some principle documents including the transcript of some relevant speeches from the While House administrations and senior security staffs. The third fragment of sources are more factual; comprising some relevant research reports, state progress reports and online archival materials.

Review, Analysis and Synthetization Process

To conduct a systematic review of the documents, I first have archived them in different aforementioned classifications. Sequentially, I reviewed the conceptual part, strategy and policy section and then finished by reviewing and incorporating the factual part.

Based on the reviewed literature I developed a research framework with the defined set of criterial for assessment of the effectiveness of big data in U.S. security sector. Prediction,

identification and mitigation of crimes and threats also decision-making were observed as major criteria for examining the effectiveness of big data utilization in U.S. security sector.

Subsequently, on the basis of the developed theoretical and research framework, the data is gathered and incorporated under specific themes. For greater consistency, at the analysis section, I tried to maintain the research framework order. Yet, the focus is on mapping the chronology of big data driven initiatives in USA and evaluation of the factual data based on strategic federal documents. Moreover, for further analysis, this study tries to posit the entire U.S big data utilization process in public and security sector in a larger enabling environment. To do so, this study has looked into three major facts such as political will, the extent of investment and collaboration with different institution across sectors.

However, the findings of this study should be understood in correspondence to its reliance on secondary data, as further triangulation of data sources could bring greater insight into this study.

CHAPTER III: LITERATURE REVIEW AND RESEARCH FRAMEWORK

Big Data in Private and Public Sector

Big data has become a buzzword both in private and public sector. Companies across industries have realized that they need to hire more data scientists. Public agencies are scrambling to put together programs to utilize big data capacity. Both government and private sector get the benefit of using big data. Governments, particularly the technologically equipped countries have initiated several projects to use big data for betterment of social services, enhancement of efficiency, transparency, effective public engagement, economic growth and national security (Gang-Hoon, Trimi & Chung 2014, 69).

About the practice of big data in both public and private sector, Gang-Hoon, Trimi and Chung, indicated that big data in government and private sector is used quite differently for different purposes. Nature of works and core objectives in these two sectors vary. Private sector aims to make more profit out of its services while the governments should serve the public; protect the nation and take care of the overall development and people's well-being. Though, for both public and private sector, big data has the same value. Companies largely use big data for designing actionable solutions such as predicting customer behavior and developing competitive edge. Similarly, governments use big data to design sustainable and long-term solutions for better management of the public institutions and public wellbeing (2014, 79. 81).

There is a huge literature, bridging big data and public sector management. Scholars such as Johann and his colleagues explain the way big data can inform the policies. Public sector as executive body of the government by relying upon reliable data, can play significant roles in leading the political changes. Initially big data initiative was led by private sector, but governments also have shown interest and put enormous efforts in using big data for better governance (Gang-Hoon, Trimi & Chung 2014, 141-150).

Höchtl et al (2016), found that big data in public sector can perform different roles to reform the administrative system. It enhances the efficiencies and ease the administrative procedures through automation of tasks. In health sector it can play a revolutionary role in detecting and diagnosis of epidemics. In areas of energy and education, management of labor market, performance management and banking, big data can play highly significant role. Similarly, it enables the government to fight against organized crimes, fraud and terrorism. (154).

Ebenezer and Spassovs' study claims that data driven-e-governance is inevitable. The advancement in technology and big data would further modernize the public sector; giving it more citizen centric approach (2018, 2). This trend is further supported by the idea of open data and open governance, as the flow of data would enable the citizens and government to interact

reciprocally. Open data in public sector, enables the citizens to oversee the government performance and bring further transparency to the procedures. On the other hand, government can gauge the public satisfaction about policies implemented by government. It is claimed that big data driven governance has taken over the Weberian and New Public Management (NPM) models (Clarke, Amanda & Margetts 2014, 393,412,413).

However, Desouza, Kevin C., and Jacob in their article challenge the overly optimistic understanding of big data by explaining the limitations and promises of big data in public sector. Authors in this study claim that dark and bright side of big data is not appropriately answered. Thus, tension exist between the promise of big data and the reality. By reviewing the literature and conducting interviews with policy makers and practitioners, they examine the potential and limitations of big data. (2017, 1052).

Though, authors have clarified the vastness of big data application area, but they focus on big data prediction power for improvement of policy outcomes. (Desouza, Kevin and & Jacob 2017, 1054-1055 & 1057-1058).

Big Data Application in Security Sector

During the recent years, the use of big data in security sector has become a common practice and flood of data has given a great opportunity for security agencies to tackle the threat in national security. Digital technologies and data analysis can and are increasingly used to identify bad actors so as to detect and prevent fraud, money laundering, bribery, terrorism, regulatory non-compliance, and other criminal activities.

A study by Benjamin (2017) illustrates the process of big data technology and analysis applications in security agencies mainly in areas of profiling, tracking and mitigating the crimes. Profiling majorly involves the biographical, biological, reputational and behavioral data which are gathered from disparate sources and analyzed for decision making. Similarly, network analysis and integration of various data sets (data fusion), Meta data, and blockchain are some techniques used by data scientists in security agencies. The study by Benjamin concludes that such technique requires a broad area of expertise and necessitates teamwork and multi-disciplinary approach. It also emphasizes on placing big data technology and its utilization within a broader strategic framework that spans beyond the simple analysis result (16-31& 39).

Similarly, a research by Chi with reference to Australian national security points out some major aspect of big data applications in security sector. Big data analytics technology has revolutionized the use of unstructured data and made it searchable and sortable. Moreover, big data has made the analysis of text, sentiments and videos conceivable. This can be done in unprecedented velocity. Furthermore, big data analytics has improved the predictive analysis. Particularly, predictive models can be a useful tool for security purposes, as it enables the security

agencies to discover highly significant information in the previous datasets and come up with some non- obvious relationships and correlation which can be used for predictive modeling (2017, 2-3).

A study conducted by Putvelde et al, draws a multidisciplinary approach to developing a common understanding of big data, also its role and limitations. This study states that big data is not always a perfect tool in security sector, as it never has the ability to replace the human power. Though they admit the usefulness of this tool to enhance the effectiveness in fighting the crimes but emphasize on maintaining big data as an integral element to human judgment. The data solely is unable to predict and effectively analyze complex situations (2017, 1399-1416).

This idea is supported by Lim and Kevin (2018) explaining the conceptual aspect of big data application in strategic intelligence. By bridging the theoretical and traditional type of intelligence, their study states that data driven intelligence, despite its advancement, is unable to replace the traditional approach. It rather can be a complementary tool to the traditional form which was qualitative in nature; relying on past historical events with inductive approach.

However, Murray and Michael explore further the challenges that US intelligence is encountered by using big data. Big data is well received by the security agencies, but its applications, issues and human capital remained challenging. On the other hand, the data utilization in security at the universities are not yet a well-established unit. There are limited number of universities providing such a highly specialized training. Despite the fact that U.S universities have overall responded to the data science demand quite promptly, as almost half of the top universities are offering data analytics and data science related courses, but they are not specialized in a sense to be applied to big data utilization in security sector (Murray, Michael 2016, 92-12).

A study by Kendrick elaborates that big data technology helps States to prevent terroristic acts where prediction plays a central. This can be done by the big data technology using big datasets; as the AI technology can draw patterns out of the huge amount of data gathered from disparate sources. In countering terrorism, it further enables the efficient sources allocated in mitigating the terroristic acts. On the other hand, it gives extraordinary ability for surveillance without being restrained and which is a risk to privacy. Therefore, concrete and clear legal frameworks are required to protect people and their rights from violations. On the other hand, this study underlines the problems of AI utilization in intelligence and security investigations. It points out the technical limitations of big data and its vulnerability to fallacies, in the absence of a well mitigated system. It raises the concern of human rights violations in use of such technology, however if data utilization is well regulated, it can be an effective tool to fighting the terrorism (McKendrick 2019, 2-33).

Policy makers and security agents understand big data differently. A study conducted by Chan, Janet and Mosses, explains the different understanding of big data, its applications, capabilities and future prospect among the security agents and policy makers. Concerning the

value of big data, security agents have underlined the data richness while the policy makers have focused on the proactivity aspect of big data. By proactivity, they refer to the capability of data in alerting the system to react well in advance prior to the crime occurrence. Despite the big data promises in enhancement of the effectiveness of law enforcement and security agencies, this study states that understanding the impact of big data in security sector require inclusive understanding. It should comprise the designers, policy makers, ultimately the broader public understanding (2017, 299-315).

However, in USA the federal agencies are gathering and storing huge amount of data encompassing the datasets, images and other types of data. In 2013, US federal agencies stored on average 1.61 petabytes of data, estimated around 2.63 petabytes in the year 2015. U.S officials have constantly underlined the importance of big data in public sector and have admitted the necessity of investment on Research and Development (R&D) (Informatica 2013, 1).

A research by Thomas A. Hodge published in March 2018 explains the approach big data enabled the Homeland Security Investigation (HSI) teams to enhance the effectiveness and efficiency of their programs in tackling the problem of human trafficking. The main objective of this study is to evaluate the capability of big data analytics in analyzing the human trafficking related data to identify the illegal network and provide a reliable base for investigation and actions. By relying on experimental data, it states that big data analytics is a useful and reliable tool for the HSI team in their decision-making process, also in fighting the crime. (2018, 12).

The main emphasis of his study is on efficiency and effectiveness of big data analytics in fighting human trafficking. To evaluate the efficiency, it compares the time or pace of manual query versus Citrus application. On the other hand, to evaluate the effectiveness it measures the number of subjects or identified network in manual versus using Citrus application in the query process (Hodge 2018, xvi). Discussing the importance of big data utilization in public sector, this study highlights the concerns over the privacy, ethical use of data and its integrity (Hodge 2018, 26).

The study by Reilly thoroughly examines NSA and U.S intelligence community; using big data to fighting terrorism. Despite the big data promises, it raises the issue of big data efficacy among the intelligence community, asking whether the use of big data is an effective tool to deterring threats. By bringing the instance of the 9/11 event, it states that the 9/11 event could be prevented if its relevant data was effectively processed and shared among the security agencies. It could have helped the US decision makers to react proactively. This study concludes that big data is an effective tool for Intelligence Community (IC) to fight terrorism and prevent the crimes. However, it states that changes have to be brought into policies of using big data by the intelligence community. It should decrease the restrains policies pose on big data utilization. Although it explicitly highlights the significance of data privacy. Additionally, by referring to NSA's case, it

emphasizes on improvement in the capacity of intelligence community to effectively process the influx of data in the digital era (Reilly 2015, 18-23).

Risks and Controversies

There is a huge controversy around big data utilization for security purposes. It can be majorly divided to ethical and technical parts.

Ethically, there is a huge risk of misuse of personal data by the governments which is widely debated in the literature. The government can simply use data for surveillance and deploying more control over the citizens (Hodge 2018, 26-27)**.** For instance, the Social Credit System (SCS) has been employed in China as constant monitoring and rating mechanism over the citizens' behavior (Bostman & Rachel 2017).

Ferguson is acutely pointing out the surveilling power of big data. His study unveils the fact that citizens can be *watched, surveilled, tracked and targeted.* Technology has outpaced the surveillance process and has given more coverage power to surveilling administration. Government is enabled to have access to purchasing choice, financial situation, political attitude, social relations and all personal preferences (2010, 7-8). Thus, the data privacy is still remaining as a major concern (Hodge & Thomas 2018, 29-30).

Technically, relying entirely on the accuracy and impartiality of big data is risky. It is a human designed product and there is a greater risk of data being influenced by the human motives. Also data can be flawed and it can have several technical problems that can usually lead to discrepancies and false decision making (Ferguson 7 & 8).

After the 9/11 event, US government embarked on a set of new initiatives to fighting crimes, particularly terrorism and targeted violence. These measures were followed by investing on big data utilization in security agencies. Yet, the literature to evaluate the effectiveness of big data utilization in correspondence to U.S policies in fighting terrorism and targeted violence is insufficient.

Research Framework

To examine the effectiveness of big data utilization in U.S security sector, different approaches can be adapted. It can be assessed on the basis of big data initiative goals, side effects, relevance, client outreach, stakeholders' involvement and even within a broader scope.

Potentially, effectiveness of big data utilization in security sector can be assessed in different stages. It can be assessed from the design perspective; examining the relevance of tools and goals by looking into the indictors of *coherence, cohesiveness and consistency*. (Bali et al. 2019,4).

To assess an initiative in public sector requires a broader approach, because the public sector itself is a multi-dimensional and complex domain. However, the concept of effectiveness in public sector is defined quite differently. In a boarder sense it is defined as an active and innovative struggle to tackle a problem. (Cohen and William 2003, 28). But in a more practical sense the effectiveness refers to comparing the intended objective of an initiative or a program with the actually achieved level of objectives (Florina 2017, 314).

Accordingly, objectives are considered as benchmark to assess the development of program within a framework (Summermatter and John. 2009, 5) because the effectiveness of a program is usually gauged based on the ration of intended and achieved results. The objectives can be broad or narrow, depending on the nature and scale of the programs. In assessing the effectiveness of an initiative, usually two major categories are taken into consideration; a) cost effectiveness b) program effectiveness. Cost effectiveness is assessed based on the "technical efficiency"; estimating the unit costs of producing well-defined outcomes. On the other hand, program effectiveness is assessed based on the measures of quality, accessibility and appropriateness. The later follows a broader perspective in assessing the effectiveness (Productivity Commission 2013, 6).

Azad Singh and Mukherjee (2019) opens a new horizon to assessing the effectiveness of an initiative. According to them an intervention effectiveness and institutional capacity are invariably interwoven elements. A successful implementation of an initiative needs a great extent of capacity that refers to *analytical proficiency*, *managerial abilities*, alignment of policy design with political environment and system (104-105). This approach is adding a highly significant value to the overall assessment approach, as it enables the researchers to measure an initiative or program in a broader context.

However, this study by following a more pragmatic and broader approach tries to assess the effectiveness of big data utilization based on the intended objectives of the big data initiatives; highlighted in the U.S official documents. The aforementioned documents and the overall debate over big data initiatives in U.S public sector are in fact the roadmap to the whole big data utilization initiative in federal and security agencies. To examine the effectiveness, this study looks for cases of improvement in security agencies particularly in crimes and threats predictions, mitigation and identification, as the result of big data utilization. It also examines the way big data usage has influenced the overall decision making in security agencies.

Moreover, for a broader analysis this study tends to examine the influence of some major environmental variables over big data utilization process. These are some broader factors including the political will, extend of investment on big data initiatives and collaboration with likeminded institutions across sectors. Political will in the context of this study refers to the intension of the U.S leadership and the white house in particular in using big data in security sector. By investment

it denotes to financial resources allocated by the U.S government to be awarded for both public and private sector in order to mobilize greater innovation and accelerate the big data utilization in security sector. Likely, the collaboration refers to the partnering strategy within U.S government to involving a wide range of likeminded institutions across sectors for the greater advancement of big data utilization.

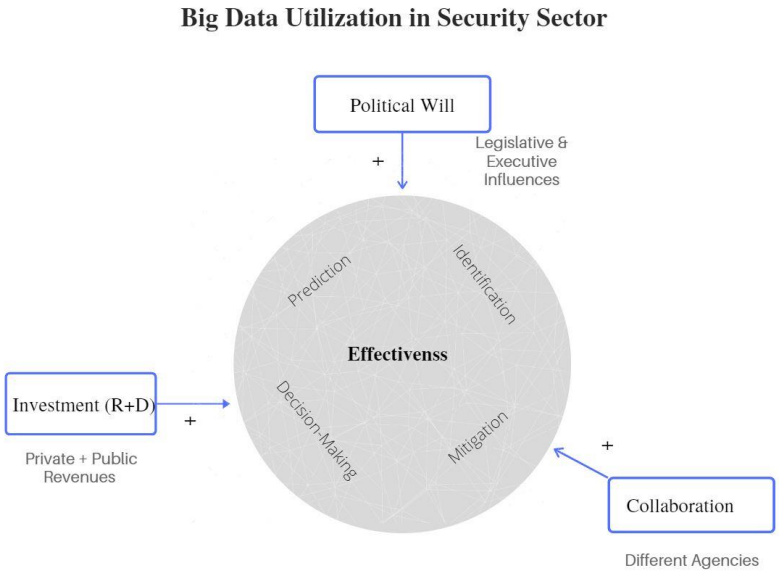

Figure I: Research Framework

To do so, this study initially evaluates the documents, policies, strategic plans and script of the U.S big data initiative in public sector, because the entire big data utilization objective in security sector is built upon those documents. Subsequently, the factual information the evidences will be evaluated in reference to the aforementioned documents to see whether the big data utilization in the real world have contributed to the quality of the performance of security agencies or not.

CHAPTER IV: RESULTS

Big Data in US Security Sector

After the 9/11 event U.S security sector experienced some drastic changes. Deployment of technology and big data applications were adapted as strategic measures for the enhancement of the effectiveness and efficiency of the institutional performance. However, the United States innovative journey toward big data utilization in public sector is explainable under two significant splits. Firstly, at the policy level, the use of big data in US security sector is intertwined with the commencement of larger debate around using big data in public sector. Secondly, at the operationalized phase, the instantaneous sparks of the big data utilization and innovation made its way out to different security agencies.

Policies and Interventions

The history of big data utilization in security sector is revolutionized in 2012 when Obama administration launched the "Big Data Research and Development Initiative". It was a comprehensive initiative to further the advancement of big data technology, enforcing the big data application in public sector. It was followed by recruitment of the first chief data scientist to the White House (2016). This initiative had a clear message for public institutions to treat data as a strategic national asset and placing national security as one of the top priorities of the U.S government in using big data (The Federal Big Data Research and Development Strategic Plan 2016, 2).

Almost all the relevant institutions were tasked to do its part in fulfilling this mission. National Science Foundation (NSF) as one of the country's leading research institution, on 4th November 2015 released a new solicitation based on which it could award around $5 million to other institutions to promote the big data application in public sector (White House Archives 2015).

Another milestone, occurred on May 23, 2016, when U.S launched "The Federal Big Data Research and Development Strategic Plan" to guide the expansion of federal initiatives and maturing the big data plans. This forged the 2012 initiative toward harnessing big data for greater analysis, information extraction and better decision making upon the large, diverse and reliable datasets. To develop the strategic plan, fifteen Federal agencies were involved under the supervision of Big Data Senior Steering Group (BD SSG) which was an interagency group within National Science and Technology Council's Networking and Information Technology Research and Development (NITRD) program. The entire idea behind the strategic plan was to promote systematically the idea of big data, its application in public sector for greater effectiveness, efficiency and better decision- making (The Federal Big Data Research and Development Strategic Plan 2016). This strategic plan functioned as a road map for all federal agencies to put forward their own agenda in using big data (2).

Department of Homeland Security (DHS) was one of the significant units to take the responsibility of deterring threats posed to the United States of America (USA**).** At the Department of Homeland Security big data utilization is guided by 2002 Homeland Security Act which authorizes the practice of data mining and use of big data analytics and technology for advancement of DHS mission. Certainly, despite the ambitious use of big data, strong emphasis and measures were put in practice to ensure the protection of the data privacy, and the entire process of big data utilization at DHS stems from three major sources (DHS Privacy Office 2010).

A) The 1974 Privacy Act
B) The 2002 E- Government Act
C) 222 section of the Homeland Security Act that emphasizes on responsible treatment of the personal data (1).

Also, DHS has its own data framework that facilitates the effective use of big data utilization and ensuring its security. To validate it further, DHS modified the framework in 2017. The framework is meant to guide the efficient and effective ways of using data for security purposes, also ensuring the protection of data privacy (Homeland Security 2017).

Additionally, the documents such as *President's Management Agenda; Leveraging Data as Strategic Asset* is another operationalized form of big data utilization guideline that orient the overall big data utilization (President's Management Agenda 2019)**.** Similarly, the institutions such as Federal Beruea of Investigation (FBI) put forward many initiatives to utilize technological innovation, particularly in gaining the benefits of big data positioning (FBI News-Testimony 2019). Almost all the guiding documents, including the ones in security agencies have a clear objective. It is to enhance the effectiveness and efficiency of the public institutions. Also, to institutionalize the evidence-based decision making.

Concern over privacy has been one of the major debates since the commencement of big data utilization, and in January 2014 President Obama assigned the *Council of Advisors on Science and Technology* (PCAST) to asses application of big data utilization in correspondence to its application over the privacy. This was intended ensure the protection and security of the public data (White House 2015).

Operationalization of Big Data in Security Areas

Following the development of strategic plans, there were several big data initiatives, emerged in public sector, including in the security agencies (Desouza, Kevin & Jacob 2017) In security sector it has transformed the policing practice, having enormous effect on operationalization of surveillance. Referring to the case of Los Angeles Police Department (LAPD), we see five major shifts in police practices (Brayne 2017).

Firstly, if we compare the shift from traditional to big data policing, risk assessment is transformed from discretionary to quantification. Operation *LASER* (Los Angeles' Strategic

Extraction and Restoration program) is one of such initiative that was funded by *Smart Policing Initiative*. LAPD was operating in a team to integrate different types of data from disparate sources; listing the "Chronic Offenders" and assigning them a number as an effort to quantify it for better prediction (Brayne 2017, 987).

Secondly, it has shifted from explanatory to predictive analytics and that is necessarily shifting toward a more proactive situation. A software developed by Pred Pol used by U.S security agencies to predict the crime occurrence. It uses three inputs such as past type of crime, place of crime and it's time to predict where future crime is more likely to occur. Based on the prediction, the police forces are deployed in the field and it is one of the fast-growing strategies (Brayne 2017, 989).

The third shift is from "Quarry to Alert based System". The large volumes of data made the alert-based system which is in contrary to the time-consuming quarry-based investigation. In alert based, the system allows the security agencies to get an alert when a specific type of targeted variable exist in the data (Brayne 2017, 990). "Lower Database Inclusion Thresholds" is another transition which heavily relies on analysis of the network of an individual through systems such as *Plantir* or *The Automatic License Plate Reader* (ALPR). For instance, a sergeant in one of his interviews described a case of a "body dump (the disposal of a dead body) that occurred in a remote location near a tourist attraction where there was an ALPR. By searching ALPR readings within the time frame that police determined the body was disposed, they captured three plates—one from Utah, one from New Mexico, and one from Compton" (992-993).

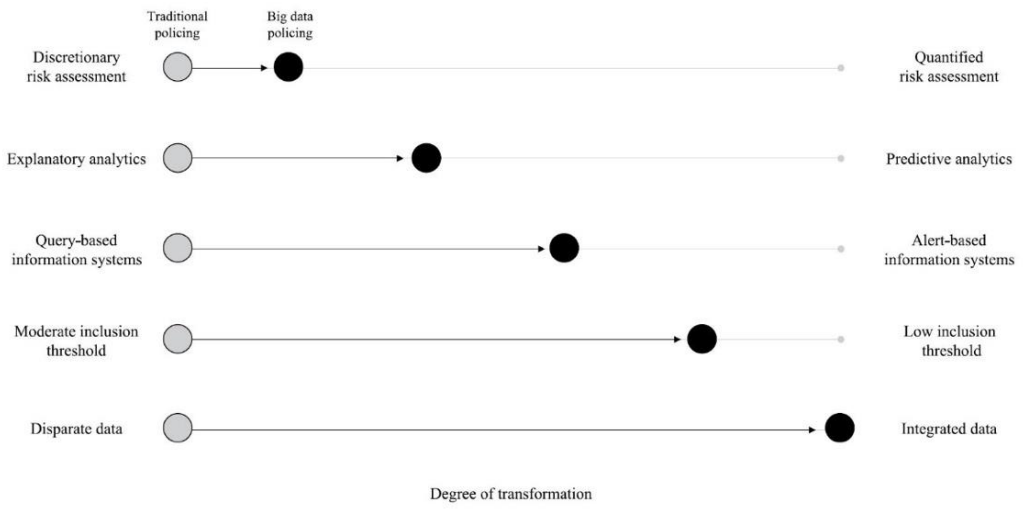

Figure II: Transformation of Police Practices toward Big Data (Brayne 2017).

According to Brayne (2017) The other revolutionary transition is the transformation of data from scattered and disparate to integrated dataset, which is now made possible by big data technology, harnessing large volumes of data from highly disparate sources.

Furthermore, FBI for its law enforcement and national security mission put forward many constructive measures to collect and use data. Christopher Wray, Director of Federal Bureau of Investigation on July 23 2019, in one of hist statement before the Senate Judiciary Committee clearly stated that FBI uses all legitimate techniques in collecting and sharing the data with its allies guided by the Joined Terrorism Task Force (JTTF). These measures were rated as effective intervention that have led to several positive outcomes in identifying and mitigating the crimes.

In some instances, such as tackling the child pornography and human trafficking, big data was used and led to surprising results. It has led to over 348 arrest of the predator of the child pornography in USA. Similarly, in the case of Las Vegas Shooting in 2017, FBI could recover one petabyte of data, as a result of newly utilized big data technology and deployment of data analysts in the field offices that has resulted in rapid and immediate response (FBI News 2019).

On the other hand, the U.S. Department of Homeland security (DHS) is a good instance of effectiveness of big data utilization in security sector. Boston Marathon bombing is one of the instances of effective use of big data in identification and mitigation of crime when the DHS by employing big data technology, could analyze around 480,000 of unstructured images for finding anomalies and certain patterns. This enabled the production of rapid result even in real time mode (Helms 2015).

Furthermore, DHS by utilizing various big data technology such as ATS-Inbound and ATS-Outbound were able to identify suspicious cargos that were associated with terrorist groups. They did it by deployment of data analytical technology which enabled them to find out useful information from the messes of data. In particular, these chunks of data could enable the officers to mitigate the risk well in advance and identify the potential cargos for physical checks (Homeland Security 2019, 10).

According to Homeland Security (2010) DARTTS is another big data technology that was realized pretty useful in identification of the criminal activity. Information gathered through (DARTTS) enabled the security agencies to identify the smuggling of US electronics to Paraguay for terroristic acts. Three individuals were arrested and persecuted. Almost all the *Data Mining Reports* by DHS from the year 2007-2018 to the congress, provide huge and strong evidences, showing large number of cases; used big data in predicting, identifying and mitigating the crimes and informing the decisions.

As the latest example, findings of the 2018 data mining report provide more instances of success and revolutionary role of big data technology in security sector. It brings up the instance

of FALACON- Road runner data analysis which helped the officers to arrest 65 criminals, 60 indictments, 18 convictions, and resulted in 35 administrative arrests. Also, it further helped the security officers in seizing great numbers of arms, ammunition, money and other vehicles equaling the value of $17,826,046, as well as 7,105 pounds of narcotics.

Besides, the application of big data enabled the security officers in identifying the high-profile launderers by compilation and analysis of the large volume of financial data (Homeland Security 2018, 44). ATLAS is another big data technology that analyzes the biographical and biometric information used for fraud detection. Only in the year 2018 it could screen over 15.5 million of combined files that could identify 2725 fraud concerns, 1458 public safety concerns and around 500 national security concerns. It alerted the security officers in advance for better crime mitigation (Homeland Security 2018, 52-53).

Another case of a jewelry robbery with the value of over 100,000 in Chevy Chase was resolved by deployment of the big data technology. FBI identified the robbers by using camera surveillance (GEOINT) in combination with the suspects' communication data (Reilly 2015, 22).

Likewise, an interesting study based on experimental data was conducted by Hodge (2018) to evaluate the effectiveness and efficiency of big data analytics in identifying the human smugglers and discovery of the criminal networks. His study finds that big data analytics has exponentially increased the effectiveness of security measures in Homeland Security Investigations.

Specifically, the Citrus application which is a big data technology used by the government has had an enormous effect on four major events. It helped the HSI team in identifying the smugglers, discovering the smugglers networks, boosting the data processing and identifying the cases for further explorations and evidence gathering (xv-xviii).

By comparing the manual and Citrus application in processing the information from the reports of the Office of Border Patrol (OBP), Hodge concludes that big data application has drastically increased the effectiveness of response to the human smuggling cases. It has not just expediated the information process in identifying the smugglers but also tells the security agents about the cases in priorities based on their seriousness.

Phone Number	Days	Manual OBP Reports	Citrus Count	Increase Percentage	Increase Average
A	**45**	9	23	155.56%	
C	**45**	8	27	237.50%	
D	**45**	12	29	141.67%	
G	**45**	8	25	212.50%	
I	**45**	6	27	350.00%	
K	**45**	9	10	11.11%	
L	**45**	8	14	75.00%	
M	**45**	6	9	50.00%	
N	**45**	5	15	200.00%	
O	**45**	7	13	85.71%	159%

Table 1: Citrus Application Effectiveness (Hodge 2018, 52)

Citrus application is proved to be a highly influential tool in using disparate communication, transaction and network data to identify and discover the smugglers. Within 45 days' time period, the study findings show a highly significant increase in processing and identifying the suspicious records (Hodge 2018, 51 & 52).

At the Department of Defense (DOD), big data utilization has led to many positive outcomes. It uses a wide variety of the big data techniques and analytical softwares. The applications of big data have effectively contributed to better decision making in various stages (Anton et al. 2019).

Environmental Factors

There are many factors that have enabled US security sector to embark on big data utilization mission. The strong political will, legislative support, huge investment and collaboration with the relevant institutions in different sectors have made USA as a front runner of big data utilization in security area.

Political Will

The idea of deployment of big data in the public sector has emerged as political agenda during Obama presidency in 2012. It commenced with a huge project "Big Data Research and Development Initiative" in which the White House was directly involved. During this time, several innovative programs were launched, different committees tasked, and recruitments were made to put this idea forward (White House 2016). Following the launch of this initiatives, most of the federal agencies begun developing their activities.

Later in December 2017, President Trump signed a new National Security Strategy to emphasize on preserving the US status as big data, technology and innovation leader globally. The strategy had a clear tone in making the AI application as the first priority in defense and security agencies (White House). The strategy encompassed a clear set of division of responsibilities; involving various institutions in the administration and public sector. Also, clear timeline and operational procedures with a defined set of tasks for every relevant agency including the allocated budget for all activities were parts of the strategy. It was well backed up by the essential legislation order (White House 2019).

Following this strategy, in 2018 DOD established the Joint Artificial Intelligence Center (JAIC) for development of the framework and advancement of AI in defense sector as its key mission. Consequently, DOD in February 2019, released its AI strategy. Similarly, in January 2019, the Director of National Intelligence (DNI) launched its initiative Augmenting Intelligence Using Machines (AIM) Initiative. This is a pretty large initiative aimed to utilizing enormous amount of data to inform relevant decisions at different levels of public institutions. In all levels of US security agencies, big data initiative was meant to enhance the effectiveness and efficiency of security programs (White House).

Investment

U.S government invested hugely for the advancement of big data application in public sector. It involved relevant stakeholders in promoting the idea of big data and AI technology in public and security sector. When the funding for Intelligence Community (IC) was disclosed, there was an enormous rise in the budget after 9/11 attack and it reached to $80.1 Billion in 2010 and it ranged between $65-70 Billion in 2015; which is still a pretty large amount (Crampton & Jeremy 2015, 23).

When Obama launched the “Big Data” initiative, six federal departments announced more than $200 million for accomplishing this mission and utilization of big data in public sector. The investments were mainly leaning toward scientific discovery, environmental and biomedical research, education and national security. Subsequently, NSF announced around $5 million in four different packages in one week to establish four regional big data innovation hubs (White House 2015).

Following the launch of initiatives on big data and AI technology, the Department of Defense (DoD) invested $250 million annually, allocating around $60 million for new projects to harness the explosion of data for better decision making at the defense department (the White House 2012). In 2018, the U.S Department of Energy organized a summit to introduce high power computers for big data utilizations. Likely, the National Science Foundation (NSF) has invested enormously to further the advancement and deployment of big data technology. In 2018, NSF grants for the science and engineering researchers to produce the supercomputer was unprecedent.

It was meant to produce the high-performance computer that can handle the highest scale and volumes of data analysis. (the White House 2020).

Additionally, many other funds allocated by NSF for turning data to information ($10 million) with University of California, Berkeley and data visualization training for undergraduates ($2 million) (the White House 2015). Also "The AIM Initiative" as strategy for augmenting and intelligence using machines was another mega program to close the gap between data and decisions in the intelligence sector (Aim Initiative).

Collaboration

To successfully promote the idea of big data utilization, the US government had to partner with large spectrum of institutions from different sectors at different levels. Academic institutions played highly significant role in taking forward the big data projects such as Big Data Initiative for Research and Development, the Defense Department's Data to Decisions program, including the Intelligence Advanced Research Projects Activity (IARPA), and the National Science Foundation's Regional Big Data Innovation Hubs (Landon & Michael 2016, 93). This idea of collaboration has been a strategic point for US government to hunt for partners not merely within states but also in private and civil society sector.

Though, the big data initiative was a joint initiative of the While House Office of Science and Technology Policy (OSTP) and various other federal agencies (White House 2015). NSF played an important role in linking more than 250 institutions including the universities from over 50 states. It established four big data regional hubs, each focusing on different challenges of big data.

- South Hub included 16 states and the District of Columbia. It was coordinated by Georgia Institute of Technology and the University of North Carolina.
- Northeast Hub, encompassed 9 states coordinated by Columbia University
- Midwest Hub, encompassing 12 sates coordinated by University of Illinois.
- West hub covered 13 states also Alaska and Hawaii. It was jointly coordinated by University of California, San Diego, the University of California, Berkeley, and the University of Washington (White House 2015).

Also, the OSTP was leading 28 cases of public- partnership initiative to harnessing big data for national priorities including the national security (White House). Furthermore, National Security Agency (NSA) for combining the cybersecurity and big data, initiated a wide array of collaboration with U.S government agencies, academia, private sector and the relevant individual researchers (White House 2015). The strategy of partnering existed even before the 2012 big data driven initiative. In 2002, the U.S government collaborated with IBM for the management of the

real time analysis of high volume of streaming data (Kim et al. 2014, 82). The U.S government has massively involved different stakeholders and sectors to innovate in this area, including the large, medium and small size institutions from academia, private sector, government, civil society and NGO sectors (White House 2016).

CHAPTER V: ANALYSIS AND CONCLUSION

An Effective Journey

Big data utilization in U.S security sector is a long-lasting tradition that reaches back to the years prior to the launch of 2012 "Big Data" initiative. U.S government had a sort of arrangement for big data utilization in some of its public sector institutions including the security agencies. But the 9/11 event was a turning point in the history of big data utilization in U.S public sector, as it unveiled the complexity of threats that could exist in the digital era even far beyond the border of the United States. Also, the changes in the technique and types of threats and crimes have posed a new challenge for the U.S security agencies. Apparently, the facts such as globalism, flow of immigrants, technological advancement and influx of data have contributed to greater comprehensions of the strategic role of technology and big data.

Agencies such as DHS has had a platform for big data utilization. Thought it was not as equipped and purposeful as it became after 2012. The 2012 intervention by Obama administration had three major messages for the entire U.S federal agencies. Firstly, it revealed to treat data as a national strategic asset. Secondly, it raised the necessity of developing a systematic approach toward using big data in public sector. Thirdly and more importantly, it invited the public agencies to a new challenge; using big data for greater innovation, effectiveness and efficiency in their performance.

Following this event, the major policies and guidelines such as "Federal Big Data Research and Development Strategic Plan" was developed. It was one of the guiding documents for big data utilization in different public sector institution including in security agencies. On the other hand, the launch of this initiative triggered a new and greater debate over the privacy of data. Scholars, analysts, policy makers and civil advocates have hugely debated this from different angles. U.S government had to think thoroughly to initiate measures to ensure the data privacy.

Thus, Obama assigned a committee to evaluate the big data utilization initiative to ensure that federal agencies are embarking on a safe mission. Following this initiative, public institutions including different security agencies such as NSA, Central Intelligence Agency (CIA), DHS, FBI and DoD developed their own projects in using big data to accomplish their own security missions.

Each agency got sufficient financial support backed up by essential legal and political provisions to invest unprecedently in the areas of big data application. Since then, several big data initiatives have been funded and implemented in the aforementioned security agencies. However, the project of big data technology went further when Trump came in office. In 2017, he released a new strategy to utilize AI technology, particularly and mainly in security agencies. It pledged more financial and political support; forging the big data and AI technology related initiatives. The

mission shifted to a smarter and more machine-driven data utilization. This intervention, however, has been enormously effective in security agencies since its beginning. It has transformed the policing and surveillance practices from traditional to big data driven. Security sector has become more proactive in a sense to having scenarios to mitigate the crimes and threats before their occurrence.

The largely existing cases and evidences prove that big data utilization in security sector has been enabling for the entire system in predicting, identifying and mitigating crimes and threats. Certainly not in a specific area, it rather has prevailed various areas such as counter terrorism, violence prevention, robbery, smuggling, fraud prevention and so on. Furthermore, their interventions and decisions are being built upon the evidences which has become an extensive practice since the emergence of big data technology and analytics.

Indicators of Success

There are some important factors that have greatly contributed to the success of this mission. Firstly, since the commencement of the big data deployment, this idea had the strong support of American leadership, as the White House and other influential governmental agencies were closely engaged. The convincing political will has played a tremendously facilitative role. Secondly, the generous budget allocation by U.S government is another vital occasion. Addition to funds allocated for public agencies, U.S government had initiated several awards to involve private sector, academia and civil society.

Through this strategy, government took further the frontiers of innovation, also used it as means to mobilize greater skills and resources for advancement of big data technology and analytics in security agencies. Though U.S government was initially confronted with limited knowledge, human capital and skills. To overcome these limitations, it has adapted a partnering strategy; linking to huge number of institutions in different sectors at national and state levels. Overall, the aforementioned factors have been extremely detrimental in expansion of the big data utilization in U.S security agencies. Currently, security institutions are prevailed by use of big data and this initiative in each of them is growing at an exponential velocity.

Lessons to Learn

This mission by U.S government has many lessons for other countries to learn. Transferring this intervention to other regions and countries with pretty similar level of advancement is a plausible alternative. Hence, considering the promises, technology and big data are offering, this study recommends the deployment of big data in security sector as a measure that transcends the effectiveness of security programs.

However, the importance of contextual realities is undeniable. It reveals us to critically evaluate the feasibility of deployment of big data in the context of fragile state and countries with poor administrative and legal systems, because utilization of big data, first and foremostly necessitates a well-established administrative infrastructure and a pro human rights legal system to ensure the data privacy. Secondly, a systematic deployment of big data in security sector, requires a locally available technology and human capital with required level of know how. Otherwise, relying on sources that exist beyond the local borders would double the financial cost.

Moreover, considering the sensitivity of security information, it is impossible to rely on human capital from abroad. More importantly, big data utilization is highly reliant on day to day to innovation. In the context of fragile state, and countries with poor knowledge hubs and underprivileged private sector, the intervention would suffer from constant dys-functionality.

More significantly, utilizing big data in security sector should remain in a proper balance between acquisition of essential information and protection of individual interests. All the security agencies, regardless of its technological and financial capabilities should remain fully obliged to the protection of human rights values such as freedom, civil liberties and privacy rights. Otherwise, in the absence of essential legal protection and aforementioned human rights values, utilization of big data in security sector would lead to digital dictatorship.

Bibliography

"A STRATEGY FOR AUGMENTING INTELLIGENCE USING MACHINES", *the Aim Initiative.*Accessed June 13, 2020, https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf

"Artificial Intelligence for the American People," *White House.* Accessed June 14, 2020, https://www.whitehouse.gov/ai/ai-american-industry/

"Big announcement big data." *Obama white house archives:* Washington: 2015. Accessed June 18, 2020.https//obamawhitehouse.archives.gov/blog/2015/11/04/big announcements-big-data

"DHS-ALL-PIA-o46 DHS Data Framework." *Homeland Security.* Accessed June 18, 2020, https://www.dhs.gov/publication/dhs-all-pia-046-b-dhs-data-framework (Accessed October, 2017).

"Executive Order on Maintaining American Leadership in Artificial Intelligence*," White Housee:* 2019. https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

"FACT SHEET: Big Data Across the Federal Government," the White House:2012**.** https://obamawhitehouse.archives.gov/the-press-office/2015/12/04/fact-sheet-big-data-across-federal-government

"Fact Sheet: PCAST Report on Big Data and Privacy: A Technological Perspective" *the White House:* 2014. Accessed June18, 2020. https://obamawhitehouse.archives.gov/the-press-office/2015/11/16/fact-sheet-pcast-report-big-data-and-privacy-technological-perspective

"FACT SHEET: The Opportunity Project – Unleashing the power of open data to build stronger ladders of opportunity for all Americans", *the White House:* 2016. https://obamawhitehouse.archives.gov/the-press-office/2016/10/06/fact-sheet-opportunity-project-unleashing-power-open-data-build.

"Leveraging Data as a Strategic Asset." *President's Management Agenda*: 2019. Accessed June 18, 2020. https://www.whitehouse.gov/wp-content/uploads/2018/03/The-President%E2%80%99s-Management-Agenda.pdf

"Oversight of the Federal Bureau of Investigation" *FBI NEWS*. Waahington:2019. Accessed June18, 2020. https://www.fbi.gov/news/testimony/oversight-of-the-federal-bureau-of-investigation-072319

Administration issues strategic plan big data research and development. Accessed June 23, 2016. https://obamawhitehouse.archives.gov/blog/2016/05/23/administration-issues-strategic-plan-big-data-research-and-development

Agbozo, Ebenezer, and Kamen Spassov. "Establishing efficient governance through data-driven e-government." In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, pp. 662-664. 2018.

Anton, Philip S., Megan McKernan, Ken Munson, James G. Kallimani, Alexis Levedahl, Irv Blickstein, Jeffrey A. Drezner, and Sydne Newberry. "Assessing the Use of Data Analytics in Department of Defense Acquisition." (2019).

B C. Dean. *The Latest Tool in Fighting Crime*. Big data.

Bali, Azad Singh, Giliberto Capano, and M. Ramesh. "Anticipating and designing for policy effectiveness." (2019): 1-13.

Big Data Senior Steering Group. "The federal big data research and development strategic plan." (2016).

Botsman, Rachel. "Big data meets Big Brother as China moves to rate its citizens." *Wired UK* 21 (2017).

Brayne, Sarah. "Big data surveillance: The case of policing." *American sociological review* 82, no. 5 (2017): 977-1008.

Chan et al., *Big Data Technology and National Security* (Australia, Methodology Report 2018).

Chan, Janet, and Lyria Bennett Moses. "Making sense of big data for security." *The British journal of criminology* 57, no. 2 (2017): 299-319.

Chi, Michael. *Big Data in National Security*. Australian Strategic Policy Institute, 2017.

Clarke, Amanda, and Helen Margetts. "Governments and citizens getting to know each other? Open, closed, and big data in public management reform." *Policy & Internet* 6, no. 4 (2014): 393-417.

Cohen, Steven, and William Eimicke. *The effective public manager: Achieving success in a changing government*. John Wiley & Sons, 2003

Crampton, JeremyW."Collec titall: National security, big data and governance." *GeoJournal* 80, no. 4 (2015): 519-531.

Data Mining Report to Congress, DHS Privacy Office. *Homeland Security*: Washington: 2010.

Desouza, Kevin C., and Benoy Jacob. "Big data in the public sector: Lessons for practitioners and scholars." *Administration & society* 49, no. 7 (2017): 1043-1064.

Desouza, Kevin C., and Benoy Jacob. "Big data in the public sector: Lessons for practitioners and scholars." *Administration & society* 49, no. 7 (2017): 1043-1064.

Ferguson, Andrew Guthrie. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: NYU Press, 2017. Accessed June 5, 2020. www.jstor.org/stable/j.ctt1pwtb27.

Florina, Popa. "Elements on the Efficiency and Effectiveness of the Public Sector." *Ovidius University Annals, Economic Sciences Series* 17, no. 2 (2017): 313-319.

Helms Josh. “Five Examples of How federal Agencies Use Big Data,” *IBM Center for The Business of Government:* 2015. http://www.businessofgovernment.org/blog/five-examples-how-federal-agencies-use-big-data.

Höchtl, Johann, Peter Parycek, and Ralph Schöllhammer. "Big data in the policy cycle: Policy decision making in the digital era." *Journal of Organizational Computing and Electronic Commerce* 26, no. 1-2 (2016): 147-169.

Hodge, Thomas A. *Application of big data analytics to support homeland security investigations targeting human smuggling networks*. Naval Postgraduate School Monterey United States, 2018.

Hodge, Thomas A. Application of big data analytics to support homeland security investigations targeting human smuggling networks. Naval Postgraduate School Monterey United States, 2018.

INFORMATICA: *Big Data for Government Drive Better Decisions for Better Policy and Program Outcomes:* Redwood:2013.

Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. "Big-data applications in the government sector." *Communications of the ACM* 57, no. 3 (2014): 78-85.

Landon-Murray, Michael. "Big data and intelligence: Applications, human capital, and education." *Journal of Strategic Security* 9, no. 2 (2016): 92-121.

Lee, In. "Big data: Dimensions, evolution, impacts, and challenges." *Business Horizons* 60, no. 3 (2017): 293-303.

Lim, Kevjn. "Big data and strategic intelligence." *Intelligence and National Security* 31, no. 4 (2016): 619-635.

McKendrick, K. "Artificial Intelligence Prediction and Counterterrorism." *London: The Royal Institute of International Affairs–Chatham House* 9 (2019).

Mukherjee, Ishani, and Azad Singh Bali. "Policy effectiveness and capacity: two sides of the design coin." *Policy Design and Practice* 2, no. 2 (2019): 103-114.

PRESS RELEASE: Obama Administration Unveils "Big Data" Initiative: Announces $200 Million in New R&D Investments," the White House: 2012. https://obamawhitehouse.archives.gov/the-press-office/2015/11/19/release-obama-administration-unveils-big-data-initiative-announces-200

Productivity Commission. "On efficiency and effectiveness: some definitions." *Staff Research Note, Canberra* (2013): 1-14.

Reilly, Brant C. "Doing More with More: The Efficacy of Big Data in the Intelligence Community." *American Intelligence Journal* 32, no. 1 (2015): 18-24.

Ritchie Hannah, Hasell Joe, Appel Cameron and Roser Max. "Terrorism." Our World in Data.org. https://ourworldindata.org/terrorism#terrorism-deaths-globally (Accessed May 30, 2020)

Sheng, Jie, Joseph Amankwah-Amoah, and Xiaojun Wang. "A multidisciplinary perspective of big data in management research." *International Journal of Production Economics* 191 (2017): 97-112.

Strang, Kenneth, David, and Zhaohao Sun. "Analyzing relationships in terrorism big data using Hadoop and statistics." *Journal of Computer Information Systems* 57, no. 1 (2017): 67-75.

Summermatter, Lukas, and John Philipp Siegel. "Defining Performance in Public Management: Variations over time and space." (2009): 34.

Van Puyvelde, Damien, Stephen Coulthart, and M. Shahriar Hossain. "Beyond the buzzword: big data and national security decision-making." *International Affairs* 93, no. 6 (2017): 1397-1416.

Van Puyvelde, Damien, Stephen Coulthart, and M. Shahriar Hossain. "Beyond the buzzword: big data and national security decision-making." *International Affairs* 93, no. 6 (2017): 1397-1416.