CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Department of Security Studies

Master's Thesis

2021 **Anastasija Šerstka**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Department of Security Studies

Big Data: A New Perspective on Conflict Resolution

Master's thesis

Author: Anastasija Šerstka

Study programme: Bezpečnostní studia (BS)

Supervisor: PhDr. Vít Střítecký, M.Phil., Ph.D.

Year of the defence: 2021

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Prague on 05.01.2021 Anastasija Šerstka

References

ŠERSTKA, Anastasija. *Big Data: A New Perspective on Conflict Resolution.* Praha, 2021. 65 pages. Master's thesis (Mgr.). Charles University, Faculty of Social Sciences, Institute of Political Studies. Department of Security Studies. Supervisor PhDr. Vít Střítecký, M.Phil., Ph.D.

Length of the thesis: 126 475 (spaces included)

Abstract

The thesis examines the role of big data in resolving modern conflicts. The study combines the concept of big data with conflict resolution theory and then applies them to three directions of conflict resolution: non-violent, violent, and conflict prevention. Each of the three groups is accompanied by a case study. This method allows a detailed understanding of various aspects related to the resolution of current conflicts using technology and big data analytics. The thesis examines empirical data associated with many innovative projects that have been implemented or are in the process of development for the resolution of ongoing conflicts – UN projects focused on big data collection, technology projects developed by the US state research centers, databases of large amounts of data related to conflicts. Based on the acquired knowledge, this work explores the big data analysis for conflict resolution, its forms, advantages, disadvantages and limitations.

Big data perspectives on the resolution of modern conflicts, based on empirical analysis, are summarized in three groups: operational (real-time data collection and processing), tactical (real-time decision-making based on big data analysis outcomes), and strategic (data-driven strategic advantage). The thesis concludes that the main advantage of big data implementation is the collection and processing (big data analysis) of large amounts of mostly unstructured conflict-related data in real-time. The results of the analysis can be applied for real-time operational and tactical decision-making, which is essential for the resolution of modern conflicts.

Abstrakt

Diplomová práce zkoumá roli velkých dat při řešení současných konfliktů. Studie kombinuje koncept velkých dat s teorií řešení konfliktů a následně aplikuje na tři směry řešení konfliktů: nenásilný, násilný a prevence konfliktu. Každá ze tří skupin je doprovázena případovou studií. Tato metoda umožňuje detailní pochopení různých aspektů, týkajících se řešení současných konfliktů pomocí technologie a analytiky velkých dat. Práce zkoumá empirická data spojená s mnoha inovativními projekty, které byly implementovány nebo jsou v procesu vývoje pro potřeby řešení současných konfliktů –

projekty OSN zaměřené na sběr velkých dat, technologické projekty vytvořené americkými státními výzkumnými centry, databáze velkých dat o konfliktech. Na základě získaných poznatků práce zkoumá analýzu velkých dat pro řešení konfliktů, její formy, výhody, nevýhody a omezení.
Perspektivy velkých dat pro řešení současných konfliktů, založené na empirické analýze, jsou shrnuty do tří skupin: operativní (sběr a zpracování dat v reálném čase), taktické (rozhodování v reálném čase na základě analýzy velkých dat) a strategické (strategická výhoda založená na implementaci výsledků datové analýzy). Práce dochází tedy k závěru, že hlavní výhodou implementace velkých dat je sběr a zpracování (analýza) velkého množství převážně nestrukturovaných dat o konfliktech v reálném čase. Výsledky analýzy pak lze použít pro rozhodování na operativní a taktické úrovni v reálném čase, což je klíčové pro řešení současných konfliktů.

Keywords

Big data, conflict resolution, big data analysis

Klíčová slova

Velká data, řešení konfliktů, analýza velkých dat

Název práce

Velká data: nová perspektiva pro řešení konfliktů

Acknowledgement

I would like to express my gratitude to my supervisor PhDr. Vít Střítecký, M.Phil., Ph.D. for his guidance and valuable advice. I would also like to extend my gratitude to my family and friends, who supported and motivated me during my studies.

Table of Contents

Introduction

Throughout history, the evolution of humankind inextricably links to the ability of people to capture, generate, and consume data. Unprocessed data are useless per se and must be processed according to the given requirement. Science has always adapted to the growth in the amount of data, and consequently, has been invented many useful technologies and techniques of data processing.

Already in the twentieth century, the quantity of data produced by human began to grow rapidly. Technological advancements led to the period of the 4th industrial revolution, when technology and data changed most realms of our lives. Along with the progress of the information society, the rapid development, and global proliferation of digital information technologies, the amount of data is growing explosively. And now, in the twenty-first century data has entitled as the “new oil”. This phrase credited to Clive Humby highlights the fact that, although data are inherently valuable, data needs processing, just as oil needs refining before its true value can be unlocked (Bridle, 2018).

Quantities of data has reached critical mass in every realm. The need to process ("refining") large amounts of data has triggered a new trend – Big Data. Initially, big data was regarded simply as a synonymous for large volumes of data. Today, the former buzzword is considered as a revolutionary concept that reaches all areas of society.

Big data relate to large sets of data that cannot be processed traditionally due to its volume. Hence, it has been necessary to create analytical tools and methods for the processing of volume data using information technologies. Currently, the main goal of big data experts is to gather information and modify algorithms for processing information through analytical software. The core element of the big data concept is datafaction (Mayer-Schönberger & Cukier, 2013) – or quantification of data. Everything we can measure should be transformed into a digital format in purpose for further analysis.

The main advantage of the current concept of big data is its universal usage. The set of technologies and techniques, big data analysis, various large data sets – all these attributes attract experts to implement big data. So, we can argue that big data technology is perspective and valuable for the future of our society. Especially for the analysis of various processes examined by social science.

The concept of big data has been already actively and successfully applied for the analysis of various processes from business to government, and medical services. The universality of the concept can serve as the motivation for the research and exploration of the role of big data in conflict resolution. In the context of a growing number of political conflicts around us, application of big data analytics for the conflict resolution process is promising. The work, thus, tries to contribute to existing scholarship on big data in conflict studies through analysis of the perspectives on conflict resolution.

This work connects two different theories: the concept of big data, which is part of information technology science, and the concept of conflict resolution as part of social sciences. The relevance of the research grows in the context of existing practice to conflict resolution and conflict prevention. Many projects and initiatives are developing to help resolve conflicts. The effectiveness of big data application for conflict resolution is primarily that successfully processed data can affect the key processes and decisions. However, the question is *how* big data affects the process of conflict resolution. Especially, if consider modern unconventional conflicts, also known as asymmetric conflicts.

Big data analysis is a potential source of valuable information that can help to understand, respond to, prevent modern conflicts. Accurate application of knowledge (as a product of complex big data analysis) can even identify targets, predict the manifestations of violence, and support development in conflict-affected states. The list of benefits of a big data application is long and heterogeneous, thus, to understand the full spectrum of perspectives the thesis *aims (1) to formulate a set of big data perspectives on conflict resolution, and (2) to find out how big data analysis influences conflict resolution and prevention considering benefits, weaknesses, and limitations.*

This work offers a qualitative analysis of the current practice of big data for resolving modern conflicts. This area is not yet much researched, but available evidence already offers valuable insights for understanding the shift from the traditional conflict resolution approach to the technology influenced one. *Therefore, the main objective of the research is to conduct a descriptive analysis that explains the role of big data in the conflict resolution process and assesses the application of big data to prevent conflict.*

Structurally, the thesis is divided into three chapters. The first, theoretical chapter, focuses on the concept of big data and theoretical basis of conflict resolution and prevention. The part that relates to big data explores the comprehensive concept of big data. The part on conflict resolution describes Peter Wallensteen's theory of conflict resolution and the mechanism of conflict prevention as a part of modern conflict resolution approach.

The second chapter examines the empirical data about big data technology implementation for conflict resolution and prevention. This chapter focuses on (1) nonviolent conflict resolution – data-driven peacekeeping, the UN big data projects; (2) violent resolution – the Internet of Military and Battlefield Things project, the US Nexus 7 project; and (3) conflict prevention – GDELT and ICEWS forecasting platforms.

The third chapter explains the essential aspects of big data analysis, as the most appropriate approach to process data collected from conflicts. Additionally, the part distinguishes big data perspectives on the resolution of modern conflicts.

Literature review

The study of the role of big data in political science is not popular yet. The lack of scholarship on big data in conflict studies became a great challenge for the research. Fortunately, there is a lot of literature that generally explains the phenomenon of big data and conflict resolution.

The main motivation for this work is the book "Big Data: A Revolution That Will Transform How We Live, Work and Think" written by analysts Viktor Mayer-Schönberger and Kenneth Cukier (2013). The piece describes how big data changes most global processes, including international relations and conflict resolution. Authors introduce a new concept – datafication, associated with big data. Also, they emphasize the role of correlation, statistical relationship, that in their opinion best characterize the essence of big data analysis.

To define the concept of big data in the thesis, the conceptualization offered by Ekbia et al. (2015), which consists of four different perspectives of big data is applicated. To investigate the characteristics of data that can be considered as big, the work uses the 7V concept described by Mazzei and Noble (2020). Big data methods and techniques are in detail described in the McKinsey Global Institute report "Big data: The next frontier for innovation, competition, and productivity"(2011). The "Big Data and Peace Security" report, written by UN Peacebuilding Support Office in collaboration with Columbia

School of International and Public Affairs (Escobal et al., 2018) is valuable source of information on big data, data analysis and conflict prevention.

To analyze the perspective of using big data on conflict resolution, the work explores the theory of conflict resolution offered by Peter Wallensteen (2012). The analysis focuses on modern conflicts, and the most relevant typology is offered by the Uppsala Conflict Data Project (UCDP, 2020), linked to the theory of Peter Wallenstein (2012). Concisely described conflict prevention theory draws on Swanström and Weissmann (2005) article.

One of the latest and most up-to-date works on conflict resolution and the role of big data in modern conflicts is the book "Small Wars, Big Data: The Information Revolution in Modern Conflict" by Eli Berman, et al. (2018). The findings from this book accompany all three parts of the work. The book has become a valuable resource of knowledge and offers: (1) definition of asymmetric and symmetric conflicts, (2) a case study on ISIS oil revenue, (3) perspectives on what big data can offer to conflict resolution.

The main sources of information for the section on big data and non-violent conflict resolution are the website and annual report (2019) of the UN Global Pulse initiative, websites of Qatalog and PulseSatellite projects. To describe data-driven peacekeeping the thesis draws knowledge from John Karlsrud article (2014).

The sub-chapter on big data and violent conflict resolution build on the role of big data technology in military interventions and actions on the battlefield. For this purpose, the work draws on critical assessment of big data technology by defense experts like Bill Holford (2017 & 2017) or Lori Cameron (2018). Also, the valuable source of information on big data technology projects are websites of IoBT REIGN and DARPA. As for Nexus 7 case study, the Sharon Weinberger's book about DARPA projects is one of few sources that describes the secret project in detail to further assess significance of Nexus 7.

The chapter about conflict prevention draws on literature about prediction models and early warning mechanisms described by the already mentioned UN and SIPA report (Escobal et al., 2018) and Emmanuel Letouzé (2013) article about the role of big data in conflict prevention and prediction. Then, the empirical part builds on Lars-Erik Cederman's and Nils B. Weidmann's article that provides an introduction to the general challenges of predicting violence and expresses skepticism about long-term prediction models. To describe big data platforms for conflict prediction are official websites of GDELT and ICEWS explored.

Methodology and Research Questions

To assess the contribution of big data for conflict resolution and conflict prevention the qualitative analysis taken in this study, combine two concepts: big data and conflict resolution. Initially, the work examines general aspects of the current concept of big data – 7V's, datafication, forms of data, data sources. Due to the lack of a coherent definition of big data, as a framework to understand big data as a concept, I choose the conceptualization of big data developed by Ekbia et al. (Ekbia et al., 2015). The concept distinct four perspectives of big data based on various problematics and agendas: product-oriented (datasets), process-oriented (data collection), cognition-oriented (data analysis), and social movement-oriented perspective (political/ social shifts enabled by big data application). These perspectives help to understand what big data means and how it this phenomenon can affect conflict resolution. Then the thesis explains the theory of conflict resolution offered by Peter Wallensteen (2012), defines asymmetric and symmetric conflicts to understand modern conflicts, and clarify conflict prevention in the context of the modern conflicts.

The second part of the thesis refers to empirical evidence on the implementation of big data for conflict resolution and prevention purposes. In this part of the work based on secondary data are described, compared, and assessed big data tools, methods, perspectives and project that aims to contribute to conflict resolution and prevention. Theoretically, this chapter opens up big data social movement perspective, identified by Ekbia et al. (2015). To systematize data, this part offers three cases based on ways and means of resolving conflicts – violent conflict resolution, non-violent conflict resolution, and conflict prevention. Each case contains a case study to help understand the significance of big data in practice.

The last, third part summarizes the findings derived from the descriptive and comparative analysis of all three cases of big data implementation. This part determines specific perspectives of big data for conflict resolution, types, strengths and limitations of big data analysis.

The main research question of the thesis is "*What is the perspective of using big data in conflict resolution?*". To find the answer to this question, the following sub-questions need to be answered:

1. Which factors influence the success of the application of big data technology in conflict resolution?

2. In which real cases have the use of big data contributed to conflict prevention or resolution?

List of Abbreviations

ACLED – Armed Conflict Location & Event Data Project

AI – artificial intelligence

BPD – barrels per day

DARPA – Defense Advanced Research Projects Agency

DoD – Department of Defense

DoDAF – Department of Defense Architechture Framework

EWRS – Early Warning and Response Systems

GCRI – Global Conflict Risk Index

GDELT – Global Database of Events, Language, and Tone

GMTI – Ground Moving Target Indicator

ICEWS – Integrated Conflict Early Warning System

IoBT REIGN – Alliance for Internet of Battlefield Things Research on Evolving Intelligent Goal-driven Networks

IoM/BT – Internet of Military / Battlefield Things

IoT – Internet of Things

ISAF – International Security Assistance Force in Afghanistan

ISIL/ISIS – Islamic State

ITU – International Telecommunication Union

MENA – Middle East and North Africa

NATO – North Atlantic Treaty Organization

NOAA – National Oceanographic and Atmospheric Association

UAV – unmanned aerial vehicle

UCDP – Uppsala Conflict Data Program

UN – United Nations

US – United States of America

1. Theoretical Background

1.1. The Concept of Big Data

1.1.1. History and definition

Although historically the concept of big data has evolved over the 20th century, the term Big Data was scientifically first used by John Mashey in 1998 (at the USENIX conference) (Mashey, 1998), however, who is the author of the term is not completely known. However, in 2008 the editor of Nature scientific journal Clifford Lynch highlighted the explosive growth of information and drew attention to the benefits and perspectives of the phenomenon of big data (Lynch, 2008). This article is considered as a starting point of the growth of popularity and interest in big data which subsequently led to the formation of the whole concept.

Although the term Big Data has become popular, there is no general agreement about the definition of the term. This is a consequence of a shared utilization of the concept between scholars, media, business and other industries (Ward & Barker, 2013). The lack of a clear definition leads to big data discourse ambiguity. This ambiguity manifests in using the term big data as a set of volume unprocessed data simultaneously with interpreting big data as technology or field of practice with a large amount of unprocessed data (processing through algorithms, special software) (Stevens & Wehrens, 2018). Therefore, for better orientation in the thesis, it is necessary to define and distinguish two different terms: *big data* as data sets or the whole concept and *big data technology* that relates to a specific technology/methodology or data processing tool. These terms, on the other hand, are inseparable and together express and form the basis of the concept of big data.

Despite the lack of a unified definition, myriad definitions which have gained some degree of attraction offered by technology corporations such as Oracle, IBM, research companies as Gartner, or information technology scholars and analysts exist. But for a better overview what does big data mean Ekbia et al. (2015) identified in the current literature four distinct conceptualizations of big data based on various problematics, and agendas that drive each perspective (Ekbia et al., 2015). The first, product-oriented perspective, focuses on the attributes of data particularly, their size, speed, structure and composition.

The second, process-oriented perspective focuses on the processes of collection, editing, and use of data. In addition, the perspective underscores the novelty of the big data processing tools. The character of these processes is typically computational, that relates to the storage, searching, data analysis. In other words, this perspective focuses on the already mentioned category of big data technology.

The next, third perspective is cognition-oriented with a focus on the human cognitive capacities and limitations that relate to data. According this perspective big data exceeds human ability to comprehend and therefore requires mediation through technological infrastructures, statistical analyses, and visualization techniques to enhance interpretability.

Fourth and the most significant is a social movement perspective that pay attention to the socioeconomic, cultural, and political shifts enabled by big data emergence. This perspective offers a suitable framework for research the role of big data in social science, including security studies, conflict management, conflict resolution, and related processes. However, the complex of perspectives offered by Ekbia et al. (2015) delineates the full scope of current big data scholarship which in their totality provide useful insights for the following research in the thesis.

1.1.2. 7 V’s of Big Data

There are many ways to think about big data, but the most popular is definition offered by Doug Laney in the Gartner's research report in 2001. He set out that "big data are high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" (Gartner, 2020). In other words, Laney enriched the concept with the “three Vs” of big data – volume, velocity, and variety. This definition was useful for its time, however, nowadays this definition is incomplete and imperfect (Mayer-Schönberger & Cukier, 2013). Nevertheless, the "3Vs" characteristics bring a valuable description of big data, which has been further extended with additional four Vs – value, viability, visualization and veracity (Stevens & Wehrens, 2018).

There are many more characteristics (>10V's), but "seven V's" as a whole are fundamental and commonly used by scholars and analysts to describe and classify big data. Three central V's: volume, velocity, and variety represent central characteristics. *Volume* represents the word "big" in big data (Mazzei & Noble, 2020) and refers to the size of data sets that need to be processed. Depending on the volume of data differ processing capabilities – the larger the data, the more complex big data technology. It means, that big data are too large in comparison with other datasets to process them with a regular computer and software. The sheer volume of data every year is growing at an incredible rate (Escobal et al., 2018, p.10). It is estimated that in 2020, internet users create 2.5 quintillion bytes of data every day (SaaS Scout, 2020).

The next element, *velocity*, refers to the speed at which the data arrives, is stored, and retrieved for processing (Mazzei & Noble, 2020). Available physical sensors, the growing number of connected devices, Internet of Things and other technology lead to increasing speed in data flow. The high velocity of data, thus, means that data analysis must adapt to the data volume. Data professionals today do not gather data over time – the faster the data can be processed, the more valuable it is (Goddart, 2019). The examples of high velocity data are Facebook messages, posts in Twitter, satellite imagery.

The third core characteristic is *variety* of data. Technological advancements allow tracking of tremendous quantity of data and make big data really "big" due to wide range of data sources. Raw, semi-structured, and unstructured data sourced from websites, social media, email, mobile data, CRM systems, IoT data, images, satellite images, sensors and many other sources produce data that can be collected, stored, processed, and analyzed (Goddart, 2019).

The next fundamental characteristic, *veracity*, relates to data quality and trustworthiness (Mazzei & Noble, 2020). High veracity data contributes in a significant way to valuable results of big data analysis. It is crucial that the data are extracted directly from the source rather than from third-party. As unreliable data sources lead to distorted results of big data analysis and bad decisions that are made on the basis of it (Mazzei & Noble, 2020).

Fifth, and essential for analysis in social studies is *viability*. It is the relevance of data that concerns the possibility of the data to be analyzed in a manner to make it relevant for decision-makers (Mazzei & Noble, 2020). Viability even uncovers the hidden relationships among significant variables, which are unrecognizable without big data analysis. This is useful for predictive analysis and decision-making process.

With similar decision-making effect is the next characteristic – *visualization*. Visualization is a visual or graphical representation of analyzed data. In other words, visualization makes information comprehensive as it is reflection of the results of big data analysis in specific visual form.

The last, seventh, and the most important characteristic is the *value* of big data. Value refers to an ability to transform amounts of data into real valuable information and knowledge (Goddart, 2019). In other words, the potential value that the data might create for business, government, healthcare and so forth (Cristobal, 2020). The interaction of data with analytical tools allows them to become significantly valuable and useful for many purposes like predictive analysis, optimization or evaluation of processes in various areas. Nevertheless, it is useless to embark on big data initiatives without proper intention or proper application and clear understanding of the value analyzed data will bring. Though the full set of characteristics of big data has still being debated, it is certain that big data are no more just a buzzword, but a new concept with extensive influence and undoubted importance for the future.

1.1.3. Big Data Processing

Increasing volume of data puts a challenge on how to collect, process and transform them into valuable information and knowledge. First and foremost, in order for any raw data to be suitable for processing, the data must be *datafy*. *Datafication*, the concept introduced by Kenneth Cukier and Victor Mayer-Schöenberger, refers to a process of putting a diverse range of data in a quantified format so it can be tabulated and analyzed (Mayer-Schönberger & Cukier, 2013). It can be even said that this is a core requirement for big data realization. Datafication intensifies as more dimensions of social life play out in digital spaces due to the rise of technology development and digitization (Southerton, 2020).

The benefit of the concept of datafication stems from the growth in the number of various sources what allows us to use and analyze information in new ways and unlock the implicit value of data. Big data deals with gigantic amounts of data that cannot be processed and analyzed using traditional techniques or tools. For the purpose of data processing, it is essential to know what forms of digital data and its sources exist and how we can collect data.

Exist three forms of raw data that have become relevant for big data processing – structured, semi-structured, and unstructured (Mazzei & Noble, 2020). *Structured* data are well-organized data that can easily be searched, collected, and processed by data mining tools because data has been already identified, organized, and classified for further use (Escobal et al., 2018, p.11). Examples of structured data are various machine-coded databases (Hammond & Weidmann, 2014), statistics, spreadsheets, barcodes. Owing to the comprehensible organization of structured data, it is easier to process them in comparison with semi-structured and unstructured data.

As for *semi-structured* data, they sit somewhere between structured and unstructured data. Such a form of data does not correspond to the organized structure of structured data but nonetheless contains tags or markers that make such data slightly more searchable (Goddart, 2019). Emails are considered as the best examples.

The third form is *unstructured* data, dispersed and unorganized with hardly unidentifiable internal structure (Escobal et al., 2018). Such datasets of various objects have no use until identified and organized properly. Unstructured data are the largest form of data as regards daily produced volume. Thus, big data are closely associated especially with unstructured data. Experts estimate that 90 percent of big data are unstructured data and the majority of big data technology designed to analyze them can handle unstructured data (Beal, 2020). Unstructured data include textual and multimedia content – various videos, images, social media posts, messages (Goddart, 2019).

Generally, structured data has been preferred by traditional systems and reporting still rely on this format of data (Naeem, 2020). However, increase in quantity of semi-structured and unstructured data sources in the past few years led to devise of big data phenomenon which shifted traditional analytics to the next level with including all three forms of data.

The evolution of data also changes the ways data are collected. For a purpose of big data analysis, data can be generally collected either actively (intentionally collected data) or passively (automatically collected data) (Escobal et al., 2018). Increasing amounts of all three forms of data, collected whether actively or passively, come from various sources. To extract valuable information, we need to distinguish various categories of big data sources. Big data due to its massive nature are not structured well, thus in order to achieve a successful outcome, it is important to know how to utilize various data sources and collect relevant data. The UN and Columbia SIPA Big Data for Peace & Security report (2018) clustered big data sources into five categories: unstructured online Information and broadcast media, physical sensors, citizen-reporting or crowdsourcing, open online datasets and data exhaust (Escobal et al., 2018, p. 12). *Unstructured online information and broadcast media* refers to data that sources from web content, information agencies, radio/TV, interactive platforms and social media like Facebook, Twitter. In other words, this is online unorganized and flexible data.

Next category, *physical sensors*, deals with satellite/infrared imagery, unmanned aerial vehicle (UAVs), security cameras, and GPS. Sensors indicate, measure, or detect changing environment and human activity such as population mobility or land use.

Citizen-reporting or crowdsourcing refers to data actively produced or submitted by citizens (Escobal et al., 2018). Data require verification but provide opportunity for feedback. Examples of citizen-reporting or crowdsourcing data are SMS crowdsourcing, OpenStreetMap, or mTrac in Uganda, that collect data on health or malaria treatment (mTrac, 2020).

The category of *open online datasets* is considered as one of the most valuable sources of big data for conflict studies. Online datasets contain already coded (machine or hand-coded) and structured data divided in various subjects (categories) for analytical purposes. The most relevant examples are World Bank Open Data, European Union Open Data Portal or U.S. Census Bureau and open datasets, that collect data about various aspects of conflicts – ACLED or the Uppsala Conflict Data Program (UCDP), ICEWS, GDELT (Escobal et al., 2018).

The last but useful type of big data sources is *data exhaust*, that comprise passively or automatically gathered data of digital activities. Digitalization gave powerful impetus to creating and monitoring networked sensors of human behavior through structured data collected for specific purposes via mobile phones, web services, transactions or information collected by governments.

In addition to the five essential categories of big data sources above-mentioned, a sixth category of the *Internet of Things (IoT)* can be added. International Telecommunication Union (ITU) has defined the concept of the Internet of Things as "a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies" (ITU, 2012, p.10). The IoT architecture consists of a network of physical sensors that automatically collect data and exchange them with specific software and hardware (electronic devices) based on the Internet. Thus, theoretically, the concept of IoT can be assigned to the category of physical sensors, but since the development of the IoT infrastructure has advanced and has begun to penetrate various areas from home to the military, it would be appropriate to classify it as a separate source of big data. Naveen Joshi, the tech expert who works on Internet of Things solutions with Big Data Analytics, claims that IoT generated data constitute a valuable independent source of big data (Joshi, 2017). The advantage of IoT lies in the ability of the sensors to provide real-time accurate information. For the big data processing purpose, the data collected by IoT may be left in its raw form or pre-processed using data mining tools or special software to be ready for analytics uses (Berman et al., 2018).

The above-mentioned process of big data collection and big data sources can be assigned to the process-oriented perspective of the concept of big data. Chronologically, the next perspective is cognition-oriented, which focuses on data analysis and infrastructure.

The definition of the concept of data analysis is as ambiguous as big data's one. However, data analysis is described as the process of data collecting, cleaning, analyzing, with final interpretation and visualization driven by a specialized computing system and software (Goddart, 2019). Big data analysis helps to analyze large volumes of all three forms of data (structured, unstructured and semi-structured) from various sources. The purpose of such analysis is to extract useful information from data for cognitive interpretation of knowledge, real-world problem solving and effective decision-making (Rao, 2018).

To carry out big data analysis, there has been developed and adapted a variety of analytical techniques and technologies, that extract and interpret valuable information. The US consulting firm McKinsey & Company in its report on big data has offered a list of big data techniques and technologies. Among the techniques are mentioned: data mining, cluster analysis, machine learning, predictive modelling, statistics, simulation, and others. To support big data techniques, technological companies developed such technologies as Hadoop, MapReduce, or R presented in the report (Manyika et al., 2011). This list is not exhaustive as researchers continue to develop new techniques and improve on existing ones in response to increasing big data demand.

1.2. Conflict Resolution

1.2.1. The concept of conflict in International Relations

During the twentieth century, the study of war and peace was as relevant as never before. After the Second World War, to explore issues of war and peace International Security Studies (ISS) has evolved as a sub-field of International Relations (Sousa, 2018, p.5). For the ISS scholars conflict studies and the concept of conflict has become crucial areas of their research.

The concept of conflict is complex, as in case with the concept of big data. There is variety of different definitions of a conflict that take into account various attributes of conflicts, but the most appropriate to this thesis is the definition of conflict offered by Peter Wallensteen (2012). He defines conflict as *a social situation in which a minimum of two actors (parties) strive to acquire at the same moment in time an available set of scarce resources* (Wallensteen, 2012). The concept, thus, consists of the three main elements of conflict: *actors, incompatibility, and action*.

Historically, conflicts waged by state actors were predominated, but contemporary conflicts are characterized by the active involvement of non-state actors. Non-state actors can be divided on either non-violent (international/regional governmental organizations, non-governmental organizations, individuals) and violent such as terrorist organizations, insurgents, rioters, private armies, individuals, and so forth. The list of examples of non-state actors in international relations is long and heterogeneous as for their nature and intentions, as such actors unlike the trained conventional forces of state actors change the very nature of modern warfare (Chenoweth & Lawrence, 2010).

The essence of any conflict encompasses disagreement(s) between at least two parties, where their demands cannot be met by the same resources at the same time. This is the way Wallensteen defines incompatibilities (Wallensteen, 2012). Different incompatibilities of conflict parties relate to four models: *Geopolitik* (conflict over territory), *Realpolitik* (conflict over political power), *Idealpolitik* (disputes over ideology), and *Kapitalpolitik* (conflicts on economic issues), which motivate parties to their actions (Wallensteen, 2012). Conflicts, however, are based not only on the general contradictions mentioned above but also on such subjective contradictions as for example religion, values, beliefs, desires, interests, etc.

If exist incompatibilities and conflicting parties, opens space for action – either violent or non-violent, depending on the rationality of the conflicting actors. The Uppsala Conflict Data Program (UCDP), the world's main conflict data provider, defines the widespread typology of conflicts if at least one of the parties in dispute is a state actor: *interstate* and *intrastate* conflicts (UCDP, 2020). Interstate conflict is a conflict between two or more states, where primary warring parties must be government parties (UCDP, 2020) with incompatibilities related to the four main categories mentioned above. Examples of interstate conflicts are World War I and World War II or in modern history invasion of Afghanistan of 2001 and the invasion of Iraq 2003.

The most diverse type of conflict in modern history is an intrastate conflict. Basically, the UCDP defines this as a conflict between a government and a non-state party, with no interference from other countries (Ibid.). But as for violent form of intrastate conflict the UCDP distinguish three different categories: *state-based armed conflicts* (between state and formally organized armed group), *non-state conflicts* (none of parties is a state) and *episodes of one-sided violence* (unilateral use of armed force against civilians, e.g., genocide, some acts of terrorism). Lots of contemporary armed conflicts, however, encompass all three different categories of violent intrastate conflicts that are located mainly in MENA region, sub-Saharan Africa and Asia (Bosetti & Einsiedel, 2015).

A conflict is a dynamic situation in which intensity level changes over a conflicts' life cycle. And just mentioned armed conflicts refer to the escalation on the highest conflict intensity level (Swanström & Weissmann, 2005), where parties on both sides resort to the use of force (Ramsbotham et al., 2011). It is not easy to define the concept of armed conflict, since it encompasses a broad spectrum of situations and characteristics. Societies involved in armed conflicts extremely suffer and pay a massive toll in loss of human life and economic, political and social disintegration (UN, 2001). Numerous armed conflicts are currently taking place around the world and according to Amnesty International by the end of 2019, 79.5 million people had been forcibly displaced worldwide, what is the largest number ever recorded (Amnesty International, 2020).

Any armed conflict, whether interstate or intrastate, can transform into state of war. The Uppsala Conflict Data Project uses quantitative approach to distinguish armed conflict from war according to the number of victims per one calendar year. Armed conflict is state-based armed conflict results in at least 25 battle-related deaths in one calendar year in comparison with war, which reaches at least 1000 battle-related deaths per calendar year (UCDP, 2020). War is one of the most destructive phenomena that human can inflict. Exists similar extreme forms of conflict as systematic repression, totalitarianism with the devastating consequences of this political system, and genocide (Wallensteen, 2012). But whether a conflict is in a non-violent or violent stage of the conflict cycle, each case generates an enormous amount of data that can be datafied through information technology and then used in analytical goals to resolve a conflict.

1.2.2. Asymmetric and symmetric conflicts

To explore modern and contemporary conflicts the concepts of symmetric and asymmetric conflict is suitable. The dichotomy predominantly refers to conflict parties, their relationship, and the conflict waging strategies. Symmetric conflicts consider conflicts of interest between relatively similar parties and capabilities as in cases of the two world wars (Berman et al., 2018). The winner of such conflict is generally the party with higher-level weapons and armies. Asymmetric conflicts, by the contrast, arise between various parties, for example between a majority and a minority, an official government and insurgents (Ramsbotham et al., 2011). Next difference is that asymmetric conflicts involve strategy of unconventional warfare, where one side enjoys advantage in technological superiority, heavy military equipment and capabilities. Asymmetric conflicts include the post-9/11 U.S. engagements in Iraq and Afghanistan, ongoing Israeli-Palestinian conflict or even continuing multi-sided civil war in Syria.

However, Eli Berman et al. (2018) in their publication "Small Wars, Big Data: The Information Revolution in Modern Conflict" emphasize the major difference between asymmetric and symmetric conflict in the context of contemporary conflicts. They describe symmetric conflict as the struggle primarily over territory, on the contrary, asymmetric conflict relates to the struggle fundamentally over people — as the people are the main source of critical information. To put it simply, Berman et al. (2018) describe asymmetric conflicts as *information-centric* (Berman et al., 2018), because success requires a flow of accurate information predominantly provided by civilians. Undoubtedly, information also plays an important role in symmetric conflict, but it is not as decisive as in asymmetric one. As holding territory is not enough to win an asymmetric conflict, the side that is able to obtain, process, and efficiently use data (e.g., if technologically prevalent party knows the location of opponent's commander, hideout, or arsenal it can remove that threat (Berman et al., 2018)), has a better chance to gain the upper hand. This fact opens up the prospects on the usage of modern information technology, especially big data, in conflict resolution of asymmetric conflicts. The empirical examples of the role of big data in asymmetric conflicts will be described in detail in the following chapters of the thesis.

1.2.3. The concept of conflict resolution

Conflict is an inherent feature of human society. Whatever the origin and nature of a conflict, it is a dynamic process that may escalate or de-escalate and is constituted by a complex interaction of attitudes and behaviors. The occurrence of conflict requires the initiation of resolution measures. The conflict resolution concept can be examined from different theoretical standpoints.

If considering the big data perspective, it will be more relevant to reflect on conflict resolution of predominantly asymmetric violent conflicts. If we study a violent form of conflict, then we can examine the conflict resolution through the concept of positive and negative peace, offered by Johan Galtung (1969). The main aim of conflict resolution, according to this perspective, is a concern to transform society in *a state of war* (violent stage of conflict) into a state of *negative peace* (the absence of violent conflict) and then transform conflict from a state of negative peace into a state of *positive peace* (the absolute absence of violent conflict and social injustice) (Sousa, 2018, p.5).

The coalescence of a conflict resolution with Galtung's concept offers a basic understanding of the conflict resolution process *per se*. Nevertheless, to define conflict resolution, we stay at the same theoretical level (as the concept of conflict), offered by Peter Wallensteen. Conflict resolution is, according to Wallensteen, *a social situation where the armed conflicting parties in a (voluntary) agreement resolve to peacefully live with – and/or dissolve – their basic incompatibilities and henceforth cease to use arms against one another* (Wallensteen, 2012). To put it simply, the process of de-escalation, cease-fire, and demilitarization is launched by both parties of a conflict. The final aim of conflict resolution is shared by both parties' perspectives regardless of process (Swanström & Weissmann, 2005).

Wallensteen's concept of the conflict resolution refers to the resolution of incompatibilities and mutual acceptance of each conflict party's existence. If interstate conflict incompatibilities refer to four categories (Geopolitik, Realpolitik, Idealpolitik, Kapitalpolitik), intrastate armed conflicts' incompatibilities, according to the UCDP, concerns government and/or territory (UCDP, 2020). The theory on conflict resolution offers seven distinct mechanisms, which help parties to dissolve their incompatibility: (1) shift parties' priorities or change of its goals, (2) division of resources as a form of compromise, (3) horse-trading, (4) shared control, (5) leave control to somebody else, (6) resorting to formal conflict resolution mechanisms, (7) leaving resolution of conflict to later (Wallensteen, 2012).

However, in armed conflicts, the resolution process and mechanisms depend on the parties and the form of the conflict. In interstate conflicts, states (conflict parties) as sovereign entities are subjects of international law, and thus in the case of a resolution, they are governed by the same or similar norms of law (international law). The situation becomes complicated in the resolution of asymmetric intrastate conflicts, as non-state violent actors have their own rules, norms, duties and thus, it is much more difficult to reach a resolution. But if a certain peace agreement is reached, such a peace may be fragile. So that the conflict does not flare up again, peacekeeping mechanisms come into play. Peacekeeping operations performed by a third party are deployed to support the implementation of a ceasefire or peace agreement (UN Peacekeeping, 2020). The mechanism of peacekeeping is a significant part of the contemporary conflict resolution process. Such multidimensional operations as a part of conflict management can facilitate the political process, protect civilians, assist in the demilitarization, demobilization, and reintegration of former combatants, protect and promote human rights and assist in restoring the rule of law (Ibid.). Peacekeeping as a widespread long-term process that can benefit from new technologies. This activity associated with conflict resolution is suitable for the collection of large amounts of data with subsequent processing and using of gained results for more effective regulation of various aspects of a post-conflict society.

1.2.4. Conflict Prevention

Peacekeeping is not the only mechanism that relates to the concept of conflict resolution. Conflict prevention, as a set of instruments used to prevent or solve disputes before they have developed into active conflicts (Swanström & Weissmann, 2005), also directly influences the resolution of conflicts. Prevention is significant before a conflict becomes violent (to prevent from any form of violence manifestation), to prevent a conflict from spreading or re-escalating, and to prevent armed conflict to escalate into full-scale war. Since, as noted above, peace following the resolution of asymmetric armed conflict tends to be unstable. In practical terms, conflict prevention interconnects with conflict resolution, peacekeeping or peacebuilding activity, as all these activities aim at peace enforcement and strengthening international and regional systems and capacities (Letouzé et al., 2013).

Traditional concept of conflict prevention consists of two categories: *direct* prevention and *structural* prevention. Direct conflict prevention refers to short-term measures that are aimed at preventing an escalation of a current or potential conflict. Structural prevention focuses on long-term measures that address the causes of a potential conflict along with potentially escalating and triggering factors (Swanström & Weissmann, 2005). Long-term prevention consists of prolonged development projects that deal with structural drivers of conflict (e.g., poverty, inequality, elite/tribe capture of the state or economy at the expense of humans) (Letouzé et al., 2013).

Depending on the conflict nature, there are different instruments for conflict prevention: preventive diplomacy or official diplomacy, social, political or political measures, etc. However, the operational level instruments of conflict prevention are the most important for the thesis purpose. Operational level includes conflict Early Warning and Response Systems (EWRS). Early warning system is defined as "a process that: (1) alerts decision makers to the potential outbreak, escalation and resurgence of violent conflict; and (2) promotes an understanding among decision makers of the nature and impacts of violent conflict" (Defontaine, 2019). Early warning involves collecting and analyzing large amounts of data on ongoing conflicts or post-conflict period, by systematically monitoring and reporting conflict indicators (Ibid.) and conflict-affected regions. Early warning systems links to response that refers to initiative that occurs when the threat of potential violent conflict is identified and that aims to manage, resolve, or prevent the violent conflict (Ibid.). Different types of response are similar to conflict management tools such as negotiation, mediation, peace-making tools, conflict preventive diplomacy, and many others. Consequently, further steps leading to conflict resolution are the competence of decision-makers.

2. Big Data and Conflict Resolution: Empirical Evidence

2.1. Conflict Resolution in the Era of Information Technology

In the twenty-first century, most of the violence we observe seems to be quite different from traditional symmetric conventional warfare. Struggles between the national states and their armies when soldiers are the main source of violence are no longer the dominant form of conflict. Instead, conflicts where collective violence by people who identify themselves as members of a certain group(s) in order to achieve political, economic, or social objectives (WHO, 2014) becomes predominant. In the thesis, this type of modern conflicts refers to an asymmetric conflict, although various authors describe them as "new wars", "small wars" (Berman et al., 2018) and many other terms. The main targets of violence in such conflicts are *civilians* whether it is the use of violence by state (against segments of people often defined in religious, ethnic, and national terms) or by non-state actors (Chenoweth & Lawrence, 2010, p. 2). Unlike conventional conflict, violent non-state actors typically seek to avoid direct fighting with state armies in asymmetric conflict as they are not as trained and equipped as a regular army.

Asymmetric intrastate armed conflict is not new phenomenon, but the reduction in the incidence of interstate conventional conflicts in contrast to increasing number of unconventional conflicts attracts scholars' and world community's attention (Gross & Gross, 2010) (Maoz & McCauley, 2008) (Gallo & Marzano, 2009) (Berman et al., 2018). Despite calls for peace, the twenty-first century is rich in asymmetric conflicts: ongoing complex Syrian Civil War and Israeli-Palestinian conflict, war in Afghanistan, or 2003 invasion of Iraq which escalated onto protracted asymmetric armed conflict.

For illustration, the continuing Israeli-Palestinian conflict between the State of Israel and Palestinian organizations such as Hamas (leading Palestinian nationalist organization), is a perfect example of structurally asymmetric state-based armed conflict (Gallo & Marzano, 2009) because: (1) it is long-termed conflict (since mid-twentieth century), (2) it is conflict between Israel powerful arms forces against Palestinian poorly equipped organizations, (3) it is armed struggle between conventional tactics and Palestinian unconventional ones such as knife attacks, suicide bombing, (4) despite long-lasting resolution process both sides have not reached successful peace agreement yet. Such long-lasting ongoing conflicts highlight the fact that existing methods, techniques, tools that would lead to conflict resolution and peace are not exhaustive. Instead, the Information Age opens new technological opportunities for conflict resolution.

The Internet, originally military network, has triggered global information revolution and quickly spread to connect various areas from academia, policy, business to individuals. (Richards & King, 2014, p. 400) The world shifted from industrial based society to upon information technology based one. Technologies spread through all areas of society and data become the essential element of today's technological world. Especially Big Data, because the volume of data is skyrocketing every year. The rapid growth of data amounts and the emergence of big data technology has modified how science is learning to make the most of data. Big data offer greater opportunities than ever before in history by a simultaneous exploration of similar and contrastive processes. Despite already mentioned lack of consensus on the big data definition, big data technology development is penetrating various areas and fields, including the security studies.

The question is, what specifically, can big data offer to security studies, especially to conflict resolution? The current literature on big data and conflict studies provides the following perspectives. First, big data allow *measuring things we never could before* (Berman et al., 2018). Each stage of conflict, whether intrastate or interstate, symmetric or asymmetric, produce enormous amounts of data scientists can datafy. If data are processed correctly, its analytical and practical capabilities of using for conflict resolution purpose is infinite. Every single activity (significant or insignificant) can be measured during a conflict – adversary's location, timeline events, routes of armed groups, civilians' migration, social media activity, etc. In other words, to resolve any conflict all possible data are essential. Especially, if we want to carry out effective multidimensional analysis of either the ongoing conflict or post-conflict events. In practice, during Iraq and Afghanistan interventions, the U.S. military was recording each activity involving its forces (including details such as the time and place of insurgent attacks, the types of attacks, outcomes (Berman et al., 2018)) with a precise time and location encoding. As for current conflicts, big data technology offers real-time analysis of real-time collected information about events in terrain (Escobal et al., 2018, p.28). Which fundamentally affects conflict decision-making. Using real-time analysis can help decision-makers (commanders, politicians) to make a decision on whether to use non-violent or violent means to eliminate the threat. And if a threat appears during peacetime, the big data real-time analysis is able to detect the threat faster than ever.

Viktor Mayer-Schönberger and Kenneth Cukier (2013), experts on big data, offer the next specific purpose of big data – *the shift from searching for causation to correlation* (Mayer-Schönberger & Cukier, 2013). Humans are conditioned to look for causation – to search the roots of things and then explain those outcomes. Nevertheless, humans' cognitive biases, limited knowledge, and experiences can often lead them down the wrong paths (Rossi, 2015). In contrast, big data provide possibilities to discover patterns and correlations in the datasets thereby offer us new and valuable insights. The correlation alerts us that something is happening rather than why something is happening (Mayer-Schönberger & Cukier, 2013). As for conflict resolution, if we do not need to know the causes of particular phenomena but determine the relationship between large volumes of unstructured data sourced from conflicts the big data analysis is thus the best approach. Furthermore, human cognitive possibilities are not even able to analyze such large quantities of data inherent in our society of the twenty-first century.

The third general perspective, according to Eli Berman et al. (2018), is that big data enable us to precisely *identify cause-and-effect relationships in ways we never could before* when analyzing conflicts (Berman et al., 2018). In comparison with previous Mayer-Schönberger's and Cukier's argument on correlation, this point may seem to be antithetical. But the cause-and-effect relationship, in this case, is not about trying to determine how each conflict works as a whole, it is rather the detailed analysis of particular fragments of conflict. As for violent armed conflicts, violence covers various aspects and depends on various things that are out of the control of researchers/analysts. Thus, a correct analysis of big data sets can recognize the effect of violence on civilians and identify where and when violent incidents happen or how violence influents development of conflict-affected region (Berman et al., 2018). Such approach is appropriate for early warning systems in purpose of conflict prevention and

peacekeeping in areas with fragile peace. The example of such analysis based on unstructured real-time data were used in Kenya during elections in anticipation of political violence in the 2010. Trained monitors at polling stations reported adverse events through mobile messaging system (SMS crowdsourcing) (Letouzé et al., 2013).
The phenomenon of big data in conflict studies has not been fully conceptualized yet. Thus, this list of big data analytical capabilities for conflict resolution offered by current big data experts can be considered incomplete, and even one that can be further expanded both horizontally and vertically. To understand the contribution of big data on conflict resolution the following sections offer a review of empirical evidence on the contemporary implementation of big data technology.

2.2. Big Data and Non-Violent Conflict Resolution

Armed conflicts considerably exhaust all conflicting parties. However, civilians suffer the most. Conflicts bring a lot of suffering and decline, whether it is the consequence of the use of violence, socio-economic problems, health or environmental problems. If the civilians' needs and problems are neglected or solved in a violent way, the conflict will be further complicated and deepened. Actually, even a cease-fire does not mean the end of a conflict. The process of solving incompatibilities is lengthy and non-violent actions, demonstrations, boycotts may only indicate that there is an interlude in the conflict (Wallensteen, 2012).
Conflict resolution mechanisms are connecting with peacekeeping and peacebuilding. Non-violent conflict resolution, thus, plays a vital role in armed conflict zones and includes peaceful means to protect humans from the power of violence. Since the main targets of violence in current asymmetric conflicts are civilians (Chenoweth & Lawrence, 2010) non-violent conflict resolution is a direct way to social transformation and justice. It can be declared that non-violent conflict resolution is a starting point of development and peace that paves a way towards widespread post-conflict prosperity.
In changing the conflict dynamics, the non-violent methods are potentially more efficient. Non-violent conflict resolution gives space to peace movements, groups, intergovernmental organizations (IGOs) or non-governmental organizations (NGOs) that work for conciliation with peaceful non-violent means (Wallensteen, 2012). There are many IGOs/NGOs that maintain international peace, but especially the United

Nations (UN) emphasize the role of new technologies and big data for non-violent conflict resolution (UN Global Pulse, 2020).

The United Nations in response to the increasing amounts of data in 2009 launched the Global Pulse initiative. The UN Global Pulse deals with big data and artificial intelligence (AI) for sustainable development, humanitarian action, and peace (UN Global Pulse, 2019). The initiative works through a network of Pulse Labs (innovation labs) which operate in Jakarta, Indonesia, Uganda, and at UN Headquarters in New York (USA). The UN Global Pulse goal is to: (1) drive exploratory research on new insights gleaned from various data sources and artificial intelligence, (2) assist UN entities, governments, development partners in making better use of data, (3) advocate for the ethical use of data and technology platforms (Ibid.).

The UN Global Pulse (2020) is based on a recognition that digital data offer opportunities to gain a better understanding of human well-being in the Information Age and to get real-time feedback on how well policy responses are working.

Global Pulse provides the UN System and public sector with tools and technical assistance needed for the adoption of an enormous amount of data, strengthen the big *data ecosystem* (data, infrastructure, visualization technology), and enables more objective decision-making (UN Global Pulse, 2020).

UN Global Pulse conducts researches and builds online tools that harness the power of big data used for global development and humanitarian action. For a non-violent conflict resolution purpose on the Global Pulse website two projects are presented: Qatalog and PulseSatellite.

Qatalog is a software analysis tool that uses speech recognition technology developed by UN Global Pulse to the detection of spoken words, pulls in public Twitter streams (Qatalog, 2020). And, most importantly for conflict resolution scholars and experts, it allows users to upload PDF documents for analysis. Qatalog allows analysts to extract useful information from sources of big data and analyze it for its purpose using a combination of optimized techniques that include translation, geolocation, and machine learning-driven text classification (Ibid.). Another advantage of this project is, that users can visualize the volumes of annotated data over time and space or can download the raw data (structured or/and unstructured) for specific analysis.

The example of successful Global Pulse big data project (forerunner of Qatalog) is Radio Content Analysis (2015) in Uganda. The project was built in collaboration with Stellenbosch University (South Africa), a center of excellence in speech recognition. The cooperation led to develop the hardware and software for an innovative radio-to-text translation platform for Uganda, that can allow the capture of public call-in-talk radio programming, translate it into text, and analyze (UN Peacekeeping, 2015). This project has sought to support Uganda, as a conflict-affected area, in incorporating the voices of Ugandan citizens into the development process (UN Global Pulse, 2016). Moreover, similar projects based on speech recognition technology can be applicable in peacekeeping to provide feedback on public information campaigns, or potentially, as an early warning mechanism (UN Peacekeeping, 2015).

The next eminent UN Global Pulse's project is PulseSatellite (since 2017). The project has been created in collaboration of UN Global Pulse and UNOSAT. PulseSatellite is web-based analysis tool that combines AI with human expertise to extract the most relevant data from satellite imagery for further use in humanitarian context (PulseSatellite, 2020). Every day, enormous amounts of satellite images are captured. In conflict areas accurate and real-time satellite image analysis is crucial for effective real-time decision-making on issues which hinder peace processes and development. In the case of the UN, analyses focus on areas of humanitarian crisis or conflict-affected zones (MENA region, Africa, Asia). For now, the beta version of PulseSatellite is being tested with a select number of UN agencies (Ibid.). According to PulseSatellite official website, the tool is used in following cases: monitoring population displacement and settlement mapping (Syria), or damage assessment (Ibid.). Analysis based on big data allows to process data faster and more efficiently. Until recently, such analysis conducted by PulseSatellite tool has been done by human analysts who spent hours counting and classifying maps, structures, and other elements (Ibid.).

Both Qatalog and PulseSatellite are the UN projects, that concern civilians' well-being as a key factor for successful conflict resolution. Peacekeeping, under the UN umbrella, is the next area of action connected with resolution of modern conflicts. As for the role of big data in the context of peacekeeping and conflict resolution, scholars often emphasize the analysis of social media activities. Especially, the Arab Spring has brought increased attention to social media (Karlsrud, 2014). The social media as a source of big data are platforms of unstructured data. The veracity of social media data in conflict-affected regions, however, may be questioned unless verified. But it is important to note that the analysis of social media in the conflict zones is significant tool. However, the aim of this thesis is to show big data from a different perspective. Mainly in terms of the data reliability. Peacekeeping and conflict resolution require more reliable and precise data than social media can offer. Furthermore, UN peace operations have always faced a problem with a lack of adequate field information (Laurence, 2020). The problem of reliability of data in the operational environment is reflected in inefficient decision-making. Thus, peacekeepers cannot protect civilians (or themselves) and conflict resolution efforts become ineffective.

The solution has become a data policy change – transition to systematic data analysis or "data-driven peacekeeping" (Ibid.). Or as John Karlsrud named it – "Peacekeeping 4.0" (Karlsrud, 2014). Data-driven peacekeeping struggles to simultaneously aggregate data about peace operations and to use data analysis for multilateral monitoring of threats and conflict zones. For example, the UN has rolled out the Situational Awareness Geospatial Enterprise (SAGE), the incident reporting situational awareness tool with a database system. The SAGE is powered by software made by well-known technology company Ushahidi that focuses on mapping and geospatial analysis (Manning, 2018). This collaboration contributes for data management and mapping during peacekeeping operations, which makes the system more reliable. The UN SAGE allows peacekeepers to record incidents like armed attacks, troop movements, abductions, and protests in Mali, Haiti, South Sudan, Lebanon, and the DRC (Manning, 2018). Currently, many reliable databases that increase the quality of conflict analysis exist. Except the UN SAGE, other important databases are World Bank Open Data, UCDP or platforms GDELT and ICEWS, which will be mention in more detail in next parts of the thesis.

Big data analytics is obviously diverse and becomes important instrument for non-violent approaches to resolve conflicts. The analysis of big data is not just about the UN peace initiatives but can also be conducted by analytics and scholars to explore various neglected aspects of particular conflict. Thus, they make a significant contribution to resolving current conflicts. The following case study is a successful example of how a proper big data analysis can disprove erroneous estimates and call into question the effectiveness of violent operational and tactical military actions.

Case Study: ISIS oil producing in Syria and Iraq

In 2017 team of researchers from the World Bank, the National Oceanographic and Atmospheric Association (NOAA), and experts from the US universities conducted the analysis to estimate the Islamic State's (also known as ISIL/ISIS or Daesh) oil revenue in Iraq and Syria. For analysis purposes, researchers initially used night-time satellite imagery measured with night-time light detecting sensors such as DMSP-OLS and SNPP-VIIIRS (Do et al., 2018).

Oil resources and oil infrastructure are significant factors influencing conflict resolution strategy in conflict-affected areas. The illicit exploitation of oil indicates unstable governance and lack of transparency (Ibid.) which substantially affects the decision-making process. The research has been highly relevant to world policy and conflict resolution in Syria, as determined the sources of ISIS financial incomes. The analysis was essential to driving ISIS out of the territory they had captured (Berman et al., 2018).

In 2013 ISIS took control of a large part of Syria and Iraq territory and then seized 42 oil production sites in both states (34 in Syria and 8 in Iraq) out of a total of 75 identified oil sites in Syria and 114 in Iraq (Do et al., 2018). Before the territory fell under ISIS control put oil output at 70,000 barrels per day (BDP) (Berman et al., 2018). During 2014 and 2015 income from oil production in areas under the Islamic State control was often cited as its largest source of income stream. Reports estimated that weekly oil revenue ranged from several million to $28 million (Do et al., 2018). Also, by estimates, after U.S. airstrikes began targeting ISIS oil facilities, the income decreased. But all estimates were based on incomplete data, and for effective conflict decision-making, it was undoubtedly important to analyze and find out the Islamic State's oil revenue and assess ISIS's long-run survival prospects had to be analyzed.

The wide disparity in the initial estimates was that they were based on data about a small number of production sites obtained at a few points in time (Berman et al., 2018). The research team completely changed the approach: they used satellite multispectral imaging (to estimate the radiant heat produced by flares at the oil fields) and ground-truth pre-war output data and effectively constructed a real-time census of oil production in areas under the Islamic State's control (Do et al., 2018). The researchers collected a large amount of structured and unstructured data and for each oil production site in the region, they used news outlets' reports, agencies' press releases, and Institute for the Study of War (ISW) maps. In the next step, they joined the satellite images data with data on oil production at ISIS-controlled and nearby oil fields obtained from IHS Energy Intelligence and Wood Mackenzie Refs (Ibid.).

The results of the research show that Islamic State performed weakly in comparison to historical trends at extracting oil in the territories of Syria and Iraq before capturing. Researchers estimated that oil production on the ISIS-controlled territory increased from circa 29,000 BPD during 2014 to an average of 40,000 BPD in 2015. Then dropped to circa 14,000 BPD in 2016, due to territory loss (Berman et al., 2018). The number of barrels alone says nothing about oil revenue. There were no reliable price data on which to calculate how to quantity (of barrels) translate into income. Some reports recounted that ISIS (as illegal terrorist group) sold oil at a discounted price – ranging from $20 to $35 dollars per barrel. Some reports stated that some fields charged $40 to $45 dollars per barrel (Do et al., 2018). The only thing is certain, the previous assumptions overestimated ISIS's oil revenue. If assume a $30 per barrel average price with no investments in maintenance, net oil revenues in 2015 would have been roughly $380M (Ibid.). The research, thus, concludes that oil revenues are well short of what needed to finance activities on the battlefield against the combined armed forces. Besides, the significant fall of global oil prices in 2014/2016 and loss of control over significant part of occupied territory in 2015 significantly aggravated the revenue fall of ISIS's oil production in 2015/2016.

The approach proposed by the team of researchers conducts a real-time census of ISIS oil production facilities with daily temporal resolution. This is a sophisticated big data analysis concerning different valuable and reliable data sources. The estimates that come from the analysis are not based on ordinary observations of one/two variables on a small number of selected locations and at a few selected dates (i.e., microdata analysis). On the contrary, the research approaches all locations in real-time analysis. The multidimensional big data analysis enables reduce bias in comparison with previous estimates of the impact of various kinds of events on oil production (Do et al., 2018). Researchers claim that the location data level production method can ensure that policy response is accurate, timely, and targeted (Ibid.). More generally, this method of data analysis can be used to planning for short-term assistance and even long-term reconstruction (post-conflict).

This case study is valuable for this thesis for several reasons. First, the big data analysis introduced by the team of researchers can help to understand economic and military activity in conflict zones. Such a method allows analyzing various types and volumes of data from different reliable sources simultaneously. And second, the method used in the case study refuted the effectiveness of violent airstrikes that targeted ISIS oil facilities. Mainly because the research completely disproved the initial estimates of ISIS oil revenues. The research outcomes can be applied to reassess strategy that can lead to conflict resolution. As for the U.S. or Russian attacks on ISIS oil facilities, there is concern over direct and long-term environmental and public health impacts. Although attacks on oil facilities (or another industrial infrastructure) are forbidden by the Geneva Conventions (Customary IHL Database, 2020), they remain common in armed conflict. The research outcomes though highlighted the possible efficiency of soft power on the resolution of the Syria civil war. The ISIS financial activity restrictions can be achieved by reducing the oil price, stopping oil smuggling, or political will from other countries to put in place stricter policies (Zwijnenburg & Waleij, 2016).

2.3. Big Data and Violent Conflict Resolution

Resolution of current asymmetric armed conflicts is a complex process. Asymmetric unconventional conflicts are about small wars, with scattered hotspots using gray zone tactics (Cohen et al., 2020), over which is difficult to take control. In spite of the fact that non-violent resolution methods are much more effective, it has to be understood that violent non-state actors are armed organized groups seek to achieve their goals by force and violence using weapons and military equipment. In this case, to achieve the reduction or even elimination of violence is difficult without confrontation on the battlefield.

After World War II, has been a priority of foreign policy in developed countries to bring order to states that have fallen into conflict (Berman et al., 2018). Since the end of the Cold War, major powers as the US, crisis management organizations like NATO, and other organizations intervene in intrastate conflicts with regulation and resolution purposes. Generally, if a dispute escalates to armed conflict or when non-violent resistance and efforts to resolve conflict fail, state military forces alone or in cooperation with organizations with military capacity can intervene in the conflict. The motives to implement a military intervention (Seybolt, 2007), however, vary depending on the causes and course of the conflict. But one thing is certain, the number of interventions will increase if the number of conflicts where traditional resolution measures fail.

The general motivation for intervention from a Western point of view is well familiar. Theoretically, the West (whether the US, the EU, NATO, and other organizations/states) as an active contributor to peace and security on the international stage focuses on containing conflict. Furthermore, in the context of asymmetric armed conflicts the aim of military interventions is to protect them from becoming a disaster, such as Syria, Somalia, or Libya cases. The escalation and collapse into civil wars in these regions have caused a large number of civilians' casualties, generated flows of migration, helped enhance international terrorism (Berman et al., 2018), and organized crime. The consequences of ongoing asymmetric armed conflicts are felt in places far from the conflict zones. For example, the refugee crisis and the growth of ISIS terrorism in Europe.

The resolution of modern asymmetric conflicts is a challenge for ongoing and future interventions. As behavioral patterns of violent non-state actors (like terrorist groups, insurgents, and others) are flexible and follow internal rules (Idler & Forest, 2015). Instead, for example, NATO's traditional pattern for intervention or organizations of collective defense are ineffective (Sauer, 2017). The use of new technologies and utilizing big data on the battlefield can become one of ways of reforming the resolution process of intrastate conflicts.

Conflicts generate an enormous amount of big data. Big data technologies can bring many advantages not only for peacekeeping or analysis of adversary's behavior but also on the battlefield. Innovation in data collection in theatre has already reshaped military practice (Berman et al., 2018), however in the realms of the defense sector, big data's impact extends beyond the bottom line (Holford, Bill, ITProPortal, 2017). Already, military applications rely on big data to drive intelligence gathering and mission-critical decision-making (Whaley, 2019). But interventions or any other missions' success directly depends on *velocity* at which the military personnel is able to collect, process, analyze, understand, and visualize quantitative data.

Big data technology on the battlefield can contribute to preciseness of strikes on targets that minimizes the use of violence and maximize the safety of civilians (Holford, Bill, PublicTechnology, 2017). New technologies based on predominantly unstructured data even decrease operational risk and increase protection of military personnel. Nowadays, a lot of attention is paid for the Internet of Military and/or Battlefield Technology (IoM/BT) – a category of Internet of things for military operations. The rise of technologies like sensors, drones and other military and airborne surveillance equipment, connected to the cohesive network (IoM/BT) produces quantities of data ranging from videos taken by drones, to text files or satellite imagery (Holford, Bill, ITProPortal, 2017). Strong edge architecture (Cameron, 2018) of IoM/BT that connects sensing, computing and communication allows military personnel to effectively respond to potential threat on the battlefield. Military benefits the IoM/BT may provide include lower costs for technology (sensors and drones), greater awareness of the combat conditions and faster real-time decision-making (Ibid.).

The Internet of Military and Battlefield Things technology is in its early stages of development. In October 2018, the U.S. Army Research Lab (ARL) awarded $25 million to the Alliance for Internet of Battlefield Things Research on Evolving Intelligent Goal-driven Networks (IoBT REIGN) to develop new predictive battlefield analytics (Ibid.). IoBT REIGN is a collaborative research alliance between the US government and university researchers for the purposes of developing a conceptual framework and empirical research on IoM/BT (IoBT REIGN, 2020). The research effort is a collaboration between ARL and Carnegie Mellon University, SRI International, University of California, Berkeley, University of California, Los Angeles University of Massachusetts, Amherst, University of Southern California, and the University of Illinois at Urbana-Champaign (UIUC) acting as the consortium lead (IoBT REIGN, 2020).

The project investigates future missions that will exploit IoM/BTs made of thousands or tens of thousands of military, adversary, and citizen nodes with a wide range of capabilities. The research investigates the use of machine intelligence and smart technology on the battlefield: from tiny occupancy sensors, drones, small on-board computer devices to powerful edge clouds with GPUs (graphics processing units) (IoBT REIGN, 2020) and other specific technology based on big data. The team of researchers is currently working on many innovative projects. An example of one of the recent projects is research on IoM/BT unconventional sensing modalities. Non-traditional sensor modality refers to physical sensors that measure some form of energy (such as sound, pressure, temperature light, RF, etc.) and process it in analytical ways (IoBT REIGN, 2020).

The team of researchers in addition to technological aspects also seeks to develop the theoretical foundations (models, and methods) and algorithms for individual tasks under various military relevant scenarios. At the industrial level of IoM/BT development are many companies engaged in the development, manufacture, sustainment of advanced technology systems. According to GlobalData (2020), a data analytics and industrial companies like Boeing, Lockheed Martin, Airbus, L3 Harris Technologies, Northrop Grumman, Thales, BAE Systems, and Leonardo DRS, will be part of the future IoM/BT revolution (GlobalData, 2019).

To understand the potential benefits of IoM/BT technology for conflict resolution, it is necessary to figure out the principle of technology operation on the battlefield. As for soldier's equipment, soldiers can wear physical sensors and computing devices connected with IoT network embedded in combat suits, helmet, weapons systems. These devices acquire a variety of static and dynamic biometrics such as a face, iris, periocular space, heart rate, gait, gestures, facial expressions (Cameron, 2018). All unstructured heterogeneous data gathered on the battlefield are stored in databases for subsequent analysis. In asymmetric conflict, it is difficult to identify enemy combatants, terrorists, insurgents as they appear as civilians (Heintschel von Heinegg, 2011). Modern physical sensors can scan biometric data to identify targets or persons who may pose a threat on the basis of the correlation of already collected data.

The key element for the optimal use of IoM/BT is artificial intelligence (AI), which allows efficient analysis of enormous amounts of unstructured data that flow at a high velocity from an increasingly large number of edge devices (sources of data) (GlobalData, 2019). While big data analytics scans through a large volume of data AI technology reduce the associated noise (Ibid.) and identify targets. Modern interventions or other military operations in civilian areas of conflict zones require modern approach to human identification and strike a target. People-centered asymmetric conflict is exactly the type of conflict where precise target identification may reduce the number of civilian casualties to a minimum, which may be one of the key factors to resolve conflict. Thus, quality of physical sensors and AI to recognize and identify targets should be a priority of IoM/BT researchers to avoid ethical dilemmas (Ibid.) during asymmetric conflicts.

Successful deployment of connected devices and unmanned systems with subsequent big data analysis will revolutionize in future interventions. It is obvious that the deployment of new data technologies and analytical skills is an inevitable part of the near future of military interventions. The U.S. Department of Defense (DoD) in order to bolster IoM/BT research in the military domain even formed some working initiatives. For example, DoD has developed the Department of Defense Architecture Framework (DoDAF) v2.0 – an informational architecture aimed at modernizing the war fighters and their infrastructure by providing guidelines on collecting, analyzing, and categorizing data (Whaley, 2019). As operations on the battlefield become more dependent on quick data analysis to make critical real-time decisions and respond to a potential threat.

The Defense Advanced Research Projects Agency (DARPA) – R&D agency of DoD – also conducts research and development on various types of technologies and analytical techniques on the battlefield. Actually, it was DARPA who created the ARPANET – the technical foundation of the Internet which launched the Information Revolution (DARPA, 2020). Now the agency is working on many data-driven technology projects.

The evidence of ongoing military implementation of new big data technologies on the asymmetric conflict's battlefield is mostly classified. It is problematic to find detailed official and verified information about the implementation of such technologies. Available empirical data on the use of big data on the battlefield of asymmetric intrastate conflicts mostly concern the analysis of the operational situation and approaches on the battlefield, and predictive analysis (Berman et al., 2018; Blair et al., 2015; Yonamine, 2013). The following case study will help to understand the essence of big data analytics on the battlefield. The case concerns not exactly successful but important implementation of project Nexus 7 – developed by DARPA – during NATO-led military mission in Afghanistan.

Case Study: DARPA Nexus 7 project

In 2010, during the International Security Assistance Force (ISAF), a NATO-led military mission in Afghanistan (2001-2014), was launched DARPA funded big data project – the Nexus 7. The project was named after a humanoid robot in the movie Blade Runner (Weinberger, 2017). The main military goal of Nexus 7 was to predict insurgency activity in Afghanistan (Ibid.) and to support ISAF operational decision-making (Berman et al., 2018) via analysis of data collected from a wide variety of systems and databases.

The Nexus 7 main researcher Peter Lee formulated a data-mining program based on the latest commercial predictive analysis but using military data from Afghanistan (Ibid.), in particular, data collected during ISAF mission. Nexus 7 was able to process all structured, semi-structured and unstructured data types in real time. The project was focused on data from various sources, from radar data and satellite imagery (Berman et al., 2018) to patterns of daily life in Afghanistan, including the costs of transportation and vegetables (Weinberger, 2017), to make prediction about insurgencies. Peter Lee claimed that Nexus 7 is a project, that tries to understand if for example the price of potatoes at local markets correlates with subsequent Taliban activity or insurgent activity (Ibid.). Satellite imagery also measured activity in rural markets in order to evaluate whether ISAF deployments were improving security situation from the civilian's point of view (Berman et al., 2018).

DARPA considered Nexus 7 as both a modern data-analysis tool and an opportunity to move beyond traditional research role into a more active analysis during wartime mission (Shachtman, 2011). The research on Nexus 7 employed the latest knowledge in quasi-experimental design, machine learning, and data mining on intelligence feeds (Weinberger, 2017) to make predictive analysis as effective as possible.

During 2010/2011 when big data only started to evolve as a concept, the main novelty of DARPA Nexus 7 project was "GMTI" (Ground Moving Target Indicator) information. In addition to monitor markets, military used data to track the travels of potential foes or to track where vehicles moved over time (Shachtman, 2011). The

GMTI had monitored what roads the locals avoided and what route they used instead. It made possible to identify Taliban's checkpoints location (Ibid.). Due to the lack of experience of big data analysis in the theatre, however, the exact knowledge that could be extracted from the GMTIs data weren't clear (Ibid.). Collected data was inconsistent. The cause of such inconsistency is that the radars and drones didn't consistently monitor over the same places at regular intervals.

The Nexus 7 was kept secret and in budget documents was described as a program that combines data analysis and prediction with social network analysis. The Pentagon characterized Nexus 7 as a successful project, but in fact there is no empirical evidence that it had any useful impact on ISAF operations (Weinberger, 2017).

It is problematic to assess the further development and success of the DARPA's project Nexus 7. There is no official mention and evidence of further implementation or upgrading of the Nexus 7 after the end of the ISAF mission in 2014. However, the case study of the Nexus 7 shows that, although the project was not long-lasting, it can be considered as the predecessor of Internet of Military and/or Battlefield Things. Theoretically, the logic of both Nexus 7 and IoM/BT is based on data collection via modern surveillance technologies and subsequent big data analysis.

The example of implementation of the Nexus 7 in Afghanistan at the time when big data technologies was just being evolve, illustrates that modern technology without methodology and proper techniques, algorithms and experts for analysis is doomed. Therefore, the IoBT REIGN team of experts, which has been actively working on IoT (IoM/BT) for military missions and defense, is expected to be more productive and successful than DARPA team.

2.4. Big Data and Conflict Prevention

In today’s world, the growing number of asymmetric conflicts is a political challenge that offers space for the implementation of modern mechanisms to prevent the emergence of new conflicts. New technologies and big data analysis can be used both for conflict resolution and for conflict prevention. The mechanism of conflict prevention aims to counteract potential triggers to widespread violent conflict (Peace Inside, 2020) or manifestations of organized violence in a fragile or conflict-affected area. As for prevention of armed conflict, the value and benefit of big data directly connect with the recognition of the international community, that conflict prevention in the Information Age can no longer be limited only to diplomatic and military initiatives (Himelfarb, 2014). In practice, conflict prevention interconnects with conflict resolution measures to enforce peace, development assistance, strengthen international and regional systems and capacities (Letouzé et al., 2013).

Conflict prevention through the use of big data is based on predicting models and systems. Technological processes to forecast emerging conflict are divided onto two categories: (a) "*prediction*" models, that aim to forecast conflict in long-term perspective (over months/years), and (b) "*early warning*" systems, which forecast conflict in the near-term perspective (Escobal et al., 2018). Big data techniques, especially machine learning can improve existing practice on conflict prevention by allowing measure and integration of multiple data types (structured, unstructured, semi-structured data) from various sources. Subsequent big data analysis then identifies correlations and patterns that point to diverse drivers of violence to forecasts conflict situation (Escobal et al., 2018).

One of the major advantages of big data are a real-time analysis of an extensive amount of data. This is the reason why models of long-term conflict prediction may be considered as inappropriate for big data analysis. Lars-Erik Cederman and Nils B. Weidmann (2017) came up with a similar argument and a high degree of skepticism about long-term models in their research on armed conflict prediction (Cederman & Weidmann, 2017). Their skepticism is based on the argument that conflicts are dynamic and complex processes powered by a huge number of factors and thus, correct conflict prediction relies on a number of limitations (Ibid.). First and foremost, the reliable outcome of any big data analysis is based on the quality and representativeness of data sets. Among other factors are contingencies (Ibid.), geopolitical changes, natural disasters, and others. For example, in 2010 it would be difficult to predict the emergence of the Russo-Ukrainian war in 2014. The analysis would be limited by wide spatial and temporal scope and inadequate data because the war was triggered after Euromaidan protests in 2013 at the national level in Ukraine. In addition, even if consider modern complex analytical techniques, for example, automated data extraction algorithms or social media signal detection which may be able to detect increased political and social tension, this does not mean that such techniques are able to predict conflict occurrence with high temporal and spatial accuracy (Ibid.). Moreover, when analyzing areas with a low probability of conflict events, long-term predictions may be erroneous.

Short-term forecasting is thus more suitable for conflict prevention than long-term prediction models. Early warning systems, as a tool for near-term forecasting, based on big data models can detect the onset of conflict from the very beginning and predict manifestations of violence and escalation of the ongoing conflict. Successful implementation of early warning systems using big data tools for conflict prevention can achieve goals at a lower cost (Escobal et al., 2018), prevent fragile regions from organized violence, and enhance political decision-making.

The contemporary scholarship on conflict early warning and response systems advocates for a more anthropo-centered approach (Letouzé et al., 2013). Humans have become the main source of large amounts of data, and people's behavior is considered an important indicator of political and social changes. Even the world community's peace efforts are now focused mainly on the development of conflict regions. The literature on early warning and conflict prevention in the context of current conflicts focuses not only on factors that trigger a conflict but even on the causes of peace in order to identify and support opportunities for prevention (Letouzé et al., 2013). These modern considerations generate opportunities for the implementation of new technology and big data for early warning systems and conflict prevention in general. As for the current early warning and response, there have been significant structural changes. These changes Emmanuel Letouzé et al. (2013) refer to the Fourth Generation of early warning and response. The main aspects that characterize the Fourth Generation of early warning are conflict areas with decentralized organizational frameworks (areas of asymmetric conflicts) and open-source technologies (software and databases for big data processing) (Letouzé et al., 2013).

Early warning systems generate a set of knowledge, based on quantitative and qualitative methods of data analysis. This should be helpful for decision-makers to formulate response – initiatives range from negotiations, peace-making, preventive diplomacy even to sanctions (Defontaine, 2019). Theoretically, early warning systems have to be linked to response instruments. Unfortunately, this point is the most problematic and limits the full-fledged big data analysis implementation. This issue will be explained in the third chapter of the thesis.

Big data mechanisms for conflict prevention closely connect with non-violent big data approaches to conflict resolution. Today, there are various UN early warning initiatives. The already mentioned Global Pulse projects and UNOSAT are tools that can be used both to monitor the conflict-affected area and to prevent conflicts in unstable and fragile regions. The SAGE, a situational awareness tool, and a database system mentioned in the previous chapter help peacekeepers identify threats. When applying machine learning techniques to SAGE data could improve the UN's approach to early warning. Because peacekeeping missions should predict violence against civilians in asymmetric conflict more accurately (Laurence, 2020).

To conduct analysis for conflict prevention purposes, the following aspects have to be considered: the temporal and spatial scope of the analysis (Cederman & Weidmann, 2017), the relevant sources and types of data, and the appropriate analytical technologies and techniques. The most relevant online datasets and analytical platforms for big data analysis on conflict prevention are especially the Global Database of Events, Language, and Tone (GDELT) and the Integrated Conflict Early Warning System (ICEWS).

The Google Jigsaw's GDELT Project, created by Kalev H. Leetaru, is an open platform that monitors the world's broadcast, print, and web news from nearly every corner of the world in over 100 languages (The GDELT Project , 2020). On the official website GDELT is described as the platform which "identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events driving our global society every second of every day, creating a free open platform for computing on the entire world (Ibid.)". The vision of the GDELT Project is to enable to provide analysis on different processes of human society, from happiness measuring through countering extremism to predict conflicts, by quantitatively codifying (datafication) of different real-time data generated by information society. GDELT uses some of the world's most sophisticated technology – powerful algorithms running on one of the most powerful server networks – that create three primary data streams with datasets of totaling trillions of datapoints, that update every 15 minutes (The GDELT Project , 2020). The GDELT database records events from 1979 to the present which allows not only analyze, but visualize, model, or examine various big datasets.

The GDELT is often compared to the Integrated Conflict Early Warning System (ICEWS). ICEWS program, developed by DARPA and now supported by Lockheed Martin, is a comprehensive system to monitor, assess, and predict national, sub-national, and international outbreaks of conflicts (ICEWS, 2020). ICEWS combines event and structured machine-coded data (Hammond & Weidmann, 2014). This platform is frequently cited example of an analytical program that is supported by various US government agencies and scholars (Escobal et al., 2018) (Hammond & Weidmann, 2014) (Ward et al., 2013).

The main purpose of ICEWS is to help political and military decision-makers by drawing their attention to a variety of events that can be relevant to conflict resolution and prevention (Escobal et al., 2018). The ICEWS platform is based on the Armed Conflict Location & Event Data Project (ACLED) database of structural indicators and tracks ongoing events through multi-language data collection (Escobal et al., 2018), identifying key locations and individuals in different context. The ICEWS project uses over 80 heterogeneous statistical and agent-based integrated models provisioned in real-time from more than 100 data sources and 250 international and regional newsfeeds to forecast crises (ICEWS, 2020). iCAST, the forecasting component of ICEWS, provides six month rolling prediction for five destabilizing Events of Interest (EOI) – Domestic Political Crisis, International Crisis, Ethnic/Religious Violence, Insurgency, and Rebellion – for 167 countries around the world (excluding the US (Ward et al., 2013)) and generates predictions with accuracy of greater than 90% (ICEWS, 2020). This great accuracy of ICEWS forecasting method of conflict prediction is based on a coalescence of various prediction models – e.g., the Bayesian statistics model or Michael D. Ward's geospatial network models (O'Brian, 2010). The coalescence of various models creates a prediction that seems to be more accurate than forecasting by one single model.

For scholars, experts, and analysts who examine crisis forecasting and factors that influence conflict resolution of ongoing intrastate conflicts, the ICEWS and GDELT (or similar databases) are valuable sources of information and knowledge. GDELT has the great benefit of being continuously updated open-source data collection, what permits its widespread use in academic and policy studies. By contrast, ICEWS is an early warning system that help US policy analysts forecast a variety of crises to which the US can respond (Ward et al., 2013). Both GDELT and ICEWS are arguably the largest data collections in social science at present. The following case study scrutinize these two repositories in the context of conflict prediction.

Case Study: ICEWS and GDELT comparative analysis

The case study examines the project of the Peace and Stability team at the European Commission's Joint Research Centre which proposes a new type of Global Conflict Risk Index (GCRI) based on ICEWS, GDELT and OEDA-Phoenix datasets. Using the original model of GCRI in combination with big data and neural networks, the team led by Matina Halkia produced a new model that monitors conflict risk in real-time following the movements and activities of violent non-state actors (terrorists, insurgents), illegal trade and organized crime syndicates (Halkia et al., 2019).

The new dynamic type of GCRI identifies each stage of the conflict development or de-escalation through conflict-related event datasets to predict conflicts through big data. To assess predictions based on empirical datasets the team of researchers adopted big data techniques and statistical methods and test them via case studies. The data on conflict-related events that sign potential triggers to violent conflicts or escalation (demonstrations, strikes, or elections-related violence) were tested on cases of Libya, Sudan, Egypt, Maldives, Nicaragua (Halkia et al., 2019).

The result of the analysis shows, that the GDELT is potentially the most accurate database to predict a conflict. The team argue, that the applied GCRI methodology on the GDELT dataset gives policymakers and experts the possibility to observe the conflict situation on a monthly basis. Even if big data based techniques seem to be able to predict a conflict in the near-time future, the analysis of the research results indicates that implementation of either GDELT or ICEWS (or OEDA-Phoenix) requires overcoming particular barriers: (1) the machine codebook algorithm is not publicly available for GDELT, which does not allow investigation on the source of potential errors, (2) the ICEWS data sources are not publicly available (only on request) so the validation of data are not easily possible, (3) common challenges need to be resolved in all datasets: false-positive rates, duplication rates, geographical/socioeconomic biases, etc. (Halkia et al., 2019).

As for the big data role in the prevention of conflicts or movements of violence, the results of this study emphasizes that big data are still evolving. The contribution of similar empirical studies for big data are that they highlight the shortcomings of datasets, models, and mechanisms for analysis. As this research demonstrates, GDELT has become a platform with the most accurate data in the context of forecasting a conflict situation on event-based data. The GDELT dataset predicted Libya's Arab Spring 2011, Sudanese protests (2018-19) and the start of Arab spring in Egypt in 2011. The ICEWS database is maybe also accurate, but due to limitations of access to many elements of data validation, the results are not precise. It can be said that the study confirmed that ICEWS is created by US developers for the purposes of American research.

This section examined empirical data regarding the use of the concept of big data for various aspects of conflict resolution. The following section will analyze the usefulness and perspectives associated with the implementation of big data for conflict resolution purposes.

3. Big Data Perspectives on Conflict Resolution

3.1.Big Data Analytics and Conflict Resolution

3.1.1. Big Data Analytics for Conflict Resolution: conditions, advantages and disadvantages

Conflict resolution is one of the key areas in social science where big data can play an important role. Conflicts generate huge amount of mostly unstructured data stem from different sources (related to nature of conflict incompatibility, conflict objectives, actors, etc.) that can be processed to generate and interpret knowledge that may lead to conflict resolution. Today, by virtue of capabilities of big data technologies every process that relates to ongoing conflicts is quantifiable and measurable. Thus, the big data analysis is the main instrument to gain the knowledge from the large quantities of data stemmed from conflicts. In order to be accurate, the analysis has to meet a number of conditions. The first is data quality and representativeness (Cederman & Weidmann, 2017) as the most important criteria for achieving accurate results that can be implemented for conflict resolution decision-making. To perform efficient analysis, data and datasets should be correspond to 7 V's of big data – be voluminous, various, reliable (veracity), viable, valuable, rapidly transmitted (velocity) and be graphically visualizable (Mazzei & Noble, 2020).

The next condition relates to relevant and reliable data sources. Analysts should be able to verify the origin and accuracy of the data. As current intrastate conflicts are considered as people-centered (Berman et al., 2018), the important conflict data source are people and their behavior. Modern conflict resolution focuses on civilians, the decreasing of violence and subsequent development, so people are considered as a very valuable source of data for generating knowledge about various aspects of conflicts. This approach, however, overlooks the fact that some regions are not connected to the internet, and some countries restrict internet access. If there is a lack of access to new technologies (computers, mobiles, sensors), internet, social networks, therefore data cannot be measured and datafied. Such a situation is observable in some conflict or violence-affected African countries – Ethiopia, Sudan or Zimbabwe. (Giles & Mway, 2020).

The following step is to select appropriate big data techniques and technologies. Powerful hardware and software, various models and algorithms that can be customized to analysis purposes are the essential attributes of big data analysis. For example, improved big data analytical models have allowed researchers to untangle the temporal and spatial dynamics of violence (Cederman & Weidmann, 2017), important for conflict prediction. Nevertheless, the adaptation and combination of various complex statistical models for the purposes of analysis are simultaneously strengths and weaknesses of big data. For example, if consider the ISIS oil revenue case study, experts had to develop their specific method based on big data models – location data level production method (Do et al., 2018). However, the variety of big data techniques, especially the wrong combination, can result in misapply with following analytical errors, which can lead to skewed results and consequently to wrong decisions. This is a problem, especially concerning the conduct of missions (peacekeeping or military interventions) and if lives of both civilians and soldiers are at stake.

The fourth condition relates to big data experts. Professional experts choose appropriate big data technology and techniques according to analysis purpose. Effective use of big data for conflict resolution or prevention requires cross-discipline collaboration (Himelfarb, 2014) between big data experts and experts on conflict resolution. However, since big data field is evolving, there is obvious deficit of experts and scholarship about implementation of big data for conflict resolution which emphasizes the relevance of this thesis.

3.1.2. Conflict Resolution and Forms of Big Data Analysis

Application of large datasets to support conflict resolution decision-making falls under the umbrella of big data analytics. The forms of big data analytics relevant to decision-making on conflict resolution can be divided on *descriptive*, *predictive*, and *diagnostic*. *Descriptive analytics* explains what happens or happened (Joshi, Naveen, 2017). This form of analysis is based on diversity of reliable databases relating to conflicts. The UN SAGE, for example, records armed attacks or troops movements in different conflict areas. New technologies such as sensors, drones or military surveillance equipment measure and collect data in real-time (Holford, Bill, PublicTechnology, 2017). Databases GDELT, ICEWS, ACLED, UCDP and similar, are also sources of various descriptive real-time data. Data obtained on the battlefield using IoM / BT technology is also subject to descriptive analysis. Descriptive analysis provides analysis of different real-time event data. Then measured events can be used for further levels of analysis to find large-scale correlations between various types of variables (events). Descriptive analysis is therefore a prerequisite for the predictive and diagnostic form of big data analytics.

The next, *predictive big data analytics*, forecasts manifestations of violence, conflict outbreaks or re-escalation through combination of big data techniques (models and algorithms). This form of analysis is applied by experts on conflict prevention and early warning to accurate short-term forecasting on the basis of platforms like GDELT or ICEWS.

The last, but the most comprehensive is *diagnostic analytics*. This type of multidimensional analysis is performed by experts and scholars who examine causation (cause-and-effect relationship), or subsequent correlations (Brinkmann, 2020), and apply theoretical knowledge on conflict resolution to interpret results. For example, the ISIS oil revenue case study. To find out whether oil production finances ISIS terrorism, a team of researchers had to analyze and correlate datasets of voluminous and various data (geospatial, economic, pre-war data, operational data about air strikes, etc.) (Do et al., 2018). And then, researchers continued to correlate outcomes with historical trends at oil extracting and the oil price on the world market. Diagnostic big data analysis is the most complex, but also the most useful form of big data analysis for a detailed exploration and assessment of the processes associated with conflict resolution.

3.2.Big Data and Conflict Resolution: Benefits and Limitations

Globally, asymmetric conflicts have become a prevalent form of conflict (Berman et al., 2018). This type of conflicts is complex and to resolve them new technologies, analytical techniques and concepts are needed. Thus, big data brings a new *data-driven conflict resolution perspective* on ongoing conflicts. The one of the main benefits of big data implementation is that big data allows to collect and analyze data in *real-time*. This makes big data analytics distinctive and different from ordinary statistics, because real-time analysis provides much more useful knowledge that can be applied by decision-makers at operational, tactical or strategical levels of the conflict. Real-time analysis is crucial for the UN peacekeeping operations, NATO's military interventions, battlefield control, or for further political decision-making. In other words, crucial for every means that aim to resolve current conflicts and overcome conflict incompatibilities. New technology and big data are even the solution for reorganizing of current interventions affected by the squeezing of operational manpower and defense budgets of nation states (Holford, Bill, ITProPortal, 2017).

The next advantage of big data is more *precise threat detection* (Berman et al., 2018) and *adversary identification* that assumes a decrease in damages and violence against civilians. Knowledge obtained through big data analysis can be applied for decisions of how to prevent violence manifestation, prevent the emergence of conflict, detect a locations or hideouts of organized groups (geospatial big data analysis), or identify persons, etc.

Information gained via big data technology for the peacekeeping operations or military interventions also brings many benefits. Big data analysis *generates the patterns, trends and outliers* for further process. Decisions based on reliable information and patterns that stems from big data can guard civilians and soldiers against the loss of life, threat of violence, and can be used for further analyses. However, there is a limitation associated with the use of data processed by new technologies concerns *data security*. Nowadays, cyberattacks or information manipulations represent growing threats to data veracity (Whaley, 2019). The essence of the risk is, that critical data can be lost or altered by third-party. If this happens, the data lose its value. This is especially dangerous for NATO or the UN missions (the UN has already been the target of offensive cyber-attacks (Laurence, 2020)), when very sensitive data are collected and processed. The leakage of important data can even have the opposite effect for conflict resolution. Knowledge gained through big data analysis can support a broad range of policy decisions and military operations (Do et al., 2018). However, the data security and reliability threat impede the widespread implementation of big data on political level.

According to Viktor Mayer-Schönberger and Kenneth Cukier (2013), the main benefit of big data is the s*hift from causation to correlation*, whatever the field of research.

However, often, as a result of processing big data are revealed correlation between events that in fact could not affect each other in any way. This phenomenon that limits big data implementation is called *spurious correlation* (Simon, 1954). The more amounts of data are analyzed, the number of spurious correlations increases. This, however, does not mean that big data analysis is ineffective. To conduct precise analysis just in addition to analysis, it is necessary to involve experts on conflict resolution, who will be able to figure out which correlations and information are of practical value and can be used to further implementation and which correlations are spurious (deviation). In addition, it opens discussion on improving big data techniques and technologies. The example is case of DARPA project Nexus 7. Although big data just began to evolve (in 2011), the setting of analysis goals led to problems with unraveling of correlations value and caused the failure of the entire program.

The next limitation of the implementation of big data concerns early warning and response systems, especially the *warning-response gap* (Letouzé et al., 2013) (Laurence, 2020). The gap reflects the fact that significant decision-making structures like the UN are not adapted to existing conflict early-warning systems. Whether it is an early warning system based on big data or not. As a result – the absence of any link between the early warning analysis and executive decision-making processes (Letouzé et al., 2013). An example of a warning-response gap is the UN peacekeepers' failure to respond when in 2016 South Sudanese soldiers launched a series of four-day lasted attacks on civilians. The attacks occurred just a mile away from the UN's headquarters, but peacekeepers did not intervene (Laurence, 2020). Therefore, unless the link between information generated by big data technology and conflict-related political decision-making is built, the implementation of big data for conflict prevention does not make sense.

However, one of the essential limitations of implementing big data for conflict resolution is *the lack of legal regulation*. To regulate big data, in most countries, regulations in the area of privacy and data are applied protection (Van Der Sloot & Van Schendel, 2016). This is insufficient given the popularity of big data and its growing use for different purposes, especially for conflict resolution. New rules at an international level should be formulated in addition to the current regulatory framework (Ibid.). For example, as far as IoM / BT and the use of related technologies are concerned, the use of big data technology should comply with the norms of both the territory where it is used and the organization (the UN, NATO) that exploits the technology.

3.3.Big data perspective on resolution of modern conflicts

Generally, based on the empirical data from previous chapters, the big data perspective on conflict resolution can be divided into three levels of big data implementation: *operational*, *tactical*, and *strategic*.

Operational perspective

Modern conflicts generate enormous amounts of measurable and quantifiable data that can be analyzed through big data technology. This is the essential argument that predetermines the whole perspective of big data on the conflict resolution process. The perspective of big data for conflict resolution at the operational level bases on *real-time collection and processing of data* (Escobal et al., 2018). Real-time data gathering and processing is the main benefit big data offer to conflict resolution. In other words, the operational perspective implies creating a network of big data sources and development of database platforms (ICEWS, GDELT, UCPD, etc.) relating to conflicts. Reliable sources are the cornerstone for any big data analysis.

The operational level primarily deals with how to make unprocessed data useful for conflict resolution. And, then how to process, adapt, understand, and exploit data. Big data analysis generates the patterns, trends, and outliers which can be used in a tactical perspective of big data on conflict resolution (Holford, Bill, PublicTechnology, 2017). Outcomes of analysis – information and knowledge – are used for further analytical or/and decision-making purposes to reach a peace agreement, that may resolve a conflict. Because the more knowledge we gain about different aspects of conflict, then the more effective tactics, and strategies that will lead to conflict resolution we can apply.

However, in order to perform an effective big data analysis, it is necessary to meet the aforementioned conditions of big data analysis and understand the appropriate purpose of the analysis. As far as peacekeeping operations are concerned, both the data selection and analysis relate mainly to data on civilians and the development of the conflict-affected area. As for the battlefield or theater, analysts are interested in data on different characteristics and movements of an adversary. In other words, an operational perspective is crucial for the subsequent successful implementation of the knowledge gained by processing data via big data technology.

Tactical perspective

The tactical level of big data perspective on conflict resolution bases on *real-time data-driven decision-making*. While operational perspective relates to data collection, processing and creating frameworks for further analysis, tactical level uses correlations, patterns, trends, analysis outliers for specific measures that lead to conflict resolution.

If consider non-violent measures to resolve conflicts, the UN has created Global Pulse initiative and platforms such as Qatalog or PulseSatellite to assist in the decision-making process of the UN development programs and measures in conflict-affected areas. As for peacekeeping, real-time decision-making during operations is also beginning to rely more and more on the power of big data technology and analysis (Karlsrud, 2014).

In theater, the modernization of soldiers' equipment is the next promising direction of the use of big data technology, especially modern physical sensors connected with IoM/BT. The new technologies based on big data gives a tactical edge. These modern devices gather enormous amounts of data, that can be rapidly transmitted into systems to drive critical real-time decisions (Whaley, 2019). Sensors can precisely detect targets (location of adversary, hideouts, routes) or identify terrorists via new technologies based on big data techniques (e.g., neural networks). The decision-maker(s) then decide what means will be used depending on the aim of the operation.

As for early warning systems, which through big data analytics can predict outbreaks of violence, the already mentioned "warning-response gap" limitation is applying to this approach. If there is a warning-response gap, then the efficiency and usefulness of big data drop to zero.

Strategic perspective

The trend of asymmetric intrastate conflict predominance will probably continue soon due to an increasingly large number of fragile countries (Berman et al., 2018). Improved weapons or military technologies employed by superpowers or organizations like the United States or NATO which come into conflict to resolve it, lead to asymmetric dominance by them over an adversary. However, violent non-state actors prefer insurgency tactics (Berman et al., 2018) rather than fight a conventional war in an attempt to meet their objectives (control of the territory or resources, etc.). In this case, the weapon supremacy of the stronger side is no longer enough to win the conflict. Therefore, in the context of current asymmetric conflicts is necessary to adjust existing strategies that will lead to conflict resolution considering the people-centric nature of ongoing asymmetric conflicts.

The challenge is the following: how to resolve asymmetric conflicts based on collective violence simultaneously ensuring the safety of civilians. The big data's contribution to conflict resolution on the strategic level bases on *data-driven strategic advantage* over an adversary (Cameron, 2018). That is, the correct application of reliable information collected by big data technology (IoM/BT devices, reliable databases) and processed through descriptive and diagnostic big data analysis can affect strategic planning and decision-making on the battlefield and beyond. This is an argument based on empirical data in the thesis. However, the successful implementation of the big data concept for the purpose of planning a strategy aimed at resolving asymmetric conflicts requires more detailed elaboration by scientists and experts of both data technology and strategic studies.

Conclusion

Big data is about large volumes of data. It is an essential condition for discussing the role of big data in conflict resolution at all. And it is precisely the condition that modern conflicts meet. Conflicts in modern history, as the research shows, are mostly asymmetric, complex, long-lasting and produce an enormous amount of data. In total, these conditions open up space for new perspectives for its solution.

Even though big data is still evolving as a concept, the technology is already beginning to penetrate the field of conflict resolution and prediction This is conditioned by the velocity information technology is spreading both at the industrial level and among people. Conflicts become information-centric, where information begins to play a strategic role. Whether intrastate, interstate or even armed conflict, it is certain that data and its processing already play and will play a significant role in the near future.

This thesis examines available empirical evidence and the current state of the implementation of big data for conflict resolution. If we compare the empirical data of all three groups of approaches associated with conflict resolution - violent (IoM/BT for strike a target), nonviolent (e.g., PulseSatellite) and prevention (prediction via EWRS), it is obvious that their common feature is data processing through big data analysis. Although big data also relates to data collection and developing databases, analytics is central to political and military decision-making that is crucial for conflict resolution.

This work scrutinizes various big data projects in the context of both state and non-state actors, whose role is growing in the context of current asymmetric conflicts. On the one hand, intergovernmental organizations such as the UN or NATO involved in research or implementation of big data conflict resolution technologies. In the case of the UN, these are development projects as Global Pulse, on the contrary, NATO has worked with the US on the military Nexus 7 program. On the other hand, non-state actors have also begun to push for research and development of big data technologies for conflict resolution, for example, Lockheed Martin Corporation and its ICEWS project, developed in collaboration with DARPA.

The main objective of the research, however, was to conduct a descriptive analysis that explains the role of big data in the conflict resolution process and assesses the application of big data to prevent conflicts. If assess the role of big data for conflict resolution, generally, big data analysis will be a central element. It means analysis as a process from the data collection through processing to the evaluation of results. Whether it is data analysis applied by experts outside the conflict area or by experts in conflict-affected zones. The advantage of big data over statistical analysis or similar basis for decision-making is not only to process huge amounts of data but the ability to process data in real-time. This is a huge benefit for decision-making that relates to conflict resolution, especially when it comes to the lives of civilians, soldiers or members of humanitarian peacekeeping missions.

A more detailed understanding of the role of big data for conflict resolution offers the following answers to research questions. The main research question is "*What is the perspective of using big data in conflict resolution?*". The answer relates to big data analysis. Conflicts generate large amounts of data, and it is important to learn how to efficiently collect, sort, encode, store, and then analyze real-time data. Big data analysis is an extremely complex process involving different methods and algorithms but obtained results can be useful for various conflict resolution and prevention real-time processes. If the analysis is performed in accordance with the objectives and meets the criteria of data reliability, then the results of the analysis can be fundamental for crucial decisions. The perspective of the use of big data includes processes such as monitoring and analyzing adversary behavior, ensuring security for civilians, developing conflict-affected areas, protecting strategic infrastructure, recording protests and attacks to conflict prediction and prevention. These findings we can summarize and classify as d*ata-driven conflict resolution.* The work also identifies three forms of big data analysis relevant to conflict resolution: descriptive, predictive (conflict prevention), and diagnostic. Each form explains different aspects of conflicts that would help identify specific knowledge important for decision-makers. To systematize different perspectives on using big data in conflict resolution the thesis distinguishes operational (real-time collection and processing data about different aspects of conflict), tactical (real-time data-driven decision-making during operations, missions, on the battlefield), and strategic perspective (data-driven strategical advantage).

The next question was as follows: *Which factors influence the success of the application of big data technology in conflict resolution?* The answer is that to perform a successful analysis, it is necessary to meet the following conditions related to big data analysis: data availability and data quality, utilization of relevant and reliable data sources, selection of appropriate big data techniques and technologies, and availability of qualified data experts. Additionally, there are a number of limitations for performing big data analysis or applying the obtained results for resolution purposes. Significant limitations include warning-response gap, shortage of experts, data security, or lack of legal regulation.

The last, third question is: *In which real cases have the use of big data contributed to conflict prevention or resolution?* Among successful big data applications for conflict resolution in conflict-affected regions are Radio Content Analysis (2015) in Uganda and violence monitoring during Kenya elections in 2010. Other successful cases of using big data for conflict resolution may include analyzes performed by experts – for example the ISIS oil revenue case study. Or ICEWS and GDELT data employment to create a new dynamic type of Global Conflict Risk Index (GCRI). However, the answer to this question remains incomplete because there is little empirical evidence of the successful application of big data in practice. Because it is either confidential data about state programs (e.g., DARPA project Nexus 7), or a simple description of various big data platforms.

It is obvious that big data for conflict resolution is just evolving. The work shows that many databases and platforms concerning data collected in conflict areas already exist, but valuable research based on database data is lacking. This data can be used to analyze various processes related to conflicts. It is the potential for future valuable scholarship on conflict resolution.

List of References

Amnesty International. (2020). *Armed Conflict | Amnesty International* . Načteno z Home Amnesty International : https://www.amnesty.org/en/what-we-do/armed-conflict/

Army Technology. (22. Jul 2020). *Internet of military things: Leading technology trends revealed* . Načteno z Army Technology | Land Defence News & Views Updated Daily: https://www.army-technology.com/comment/internet-of-military-things-leading-technology-trends-revealed/

Beal, V. (2020). *What is Unstructured Data? | Webopedia* . Načteno z Webopedia: Tech Dictionary for Students, Educators & IT Professionals : https://www.webopedia.com/definitions/unstructured-data/

Berman, E., Felter, J. H., Shapiro, J. N., & McIntyre, V. (2018). *Small wars, big data: The information revolution in modern conflict*. ISBN 9780691177076

Blair, R., Blattman, C., & Hartman, A. (2015). Predicting local violence. Available at SSRN 2497153.

Bosetti, Louise, and Sebastian Einsiedel. 2015. *Intra-State Based Armed Conflicts: Overview of Global and Regional Trends (1990–2013)*. Tokyo: United Nations University Centre for Policy Research.

Bridle, J. (2018). *Opinion: Data isn't the new oil — it's the new nuclear power* | . Načteno z TED: Ideas worth spreading : https://ideas.ted.com/opinion-data-isnt-the-new-oil-its-the-new-nuclear-power/

Brinkmann, B. (2020). *Comparing Descriptive, Predictive, Prescriptive, and Diagnostic Analytics* . Načteno z Homepage | Logi Analytics : https://www.logianalytics.com/predictive-analytics/comparing-descriptive-predictive-prescriptive-and-diagnostic-analytics/

Cameron, L. (2018). *IoT Meets the Military | IEEE Computer Society* . Načteno z IEEE Computer Society : https://www.computer.org/publications/tech-news/research/internet-of-military-battlefield-things-iomt-iobt

Cederman, Lars-Erik a Nils B. Weidmann. Predicting armed conflict: Time to adjust our expectations? Science [online]. 2017, 355(6324), 474-476. ISSN 00368075. doi:10.1126/science.aal4483

Chenoweth Erica and Lawrence Adria. Rethinking Violence: States and Non-StateActors in Conflict. Cambridge, MA: MIT Press, 2010

Cohen, Raphael S., Nathan Chandler, Shira Efron, Bryan Frederick, Eugeniu Han, Kurt Klein, Forrest E. Morgan, Ashley L. Rhoades, Howard J. Shatz, and Yuliya Shokh, Peering into the Crystal Ball: Holistically Assessing the Future of Warfare. Santa Monica, CA: RAND Corporation, 2020. https://www.rand.org/pubs/research_briefs/RB10073.html.

Cristobal, S. (17. Jun 2020). *Two more V's in Big Data: Veracity and Value - Datascience.aero* . Načteno z Home - Datascience.aero : https://datascience.aero/big-data-veracity-value/

Customary IHL Database. (2020). *Customary IHL - Practice Relating to Rule 43. Application of General Principles on the Conduct of Hostilities to the Natural Environment* . Načteno z International Committee of the Red Cross : https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2_rul_rule43

DARPA. (2020). *ARPANET*. Načteno z Defense Advanced Research Projects Agency : https://www.darpa.mil/about-us/timeline/arpanet

Defontaine, C. (15. Feb 2019). *Setting up early warning and response systems to prevent violent conflicts and save lives* . Načteno z World Bank Blogs : https://blogs.worldbank.org/dev4peace/setting-early-warning-and-response-systems-prevent-violent-conflicts-and-save-lives

Do, Q. T., Shapiro, J. N., Elvidge, C. D., Abdel-Jelil, M., Ahn, D. P., Baugh, K., & Bazilian, M. D. (2018). Terrorism, geopolitics, and oil security: Using remote sensing to estimate oil production of the Islamic State. *Energy research & social science*, *44*, 411-418. doi:10.1016/j.erss.2018.03.013

Ekbia, Hamid & Mattioli, Michael & Kouper, Inna & Arave, G. & Ghazi, Ali & Bowman, Timothy & Suri, Venkata & Tsou, Andrew & Weingart, Scott & Sugimoto, Cassidy. (2015). Big Data, Bigger Dilemmas: A Critical Review. *Journal of the Association for Information Science and Technology*. 66. 10.1002/asi.23294.

Escobal, L., Finnegan, K., Kim, H., Park, H., Schalkwyk, J., & Zlotnik, A. (2018). *Big Data for Peace and Security.* Načteno z Columbia SIPA | | School of International and Public Affairs : https://www.sipa.columbia.edu/academics/capstone-projects/using-big-data-prevent-conflict

Gallo, G., & Marzano, A. (2009). The Dynamics of Asymmetric Conflicts: The Israeli-Palestinian Case. Journal of Conflict Studies, 29. Retrieved from https://journals.lib.unb.ca/index.php/JCS/article/view/15231

Galtung, J. (1969). Violence, peace and peace research. Journal of peace research, 6(3), 167-191.

Gartner. (2020). *Definition of Big Data - Gartner Information Technology Glossary*. Načteno z Gartner: https://www.gartner.com/en/information-technology/glossary/big-data

Giles, C., & Mway, P. (02. Nov 2020). *Africa internet: Where and how are governments blocking it? - BBC News*. Načteno z BBC - Homepage: https://www.bbc.com/news/world-africa-47734843

GlobalData. (2019). *Internet of Military Things - Thematic Research* . Načteno z GlobalData Report Store : https://store.globaldata.com/report/gddef-tr-s007--internet-of-military-things-thematic-research/

Goddart, W. (12. Jul 2019). *Big Data Core Characteristics - Discover the core characteristics*. Načteno z Home - ITChronicles : https://itchronicles.com/big-data/what-are-the-core-characteristics-of-big-data

Gross, M., & Gross, M. L. (2010). Moral dilemmas of modern war: torture, assassination, and blackmail in an age of asymmetric conflict. Cambridge University Press.

Halkia, S., Ferri, S., Papazoglou, M., Van Damme, M., Jenkinson, G., Baumann, K. and Thomakos, D., Dynamic Global Conflict Risk Index, EUR 30011 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-14133-4 (online), doi:10.2760/846412 (online), JRC118701.

Hammond, J. a N.B. Weidmann. Using machine-coded event data for the micro-level study of political violence. Research and Politics [online]. 2014, 1(2). ISSN 20531680. doi:10.1177/2053168014539924

Heintschel von Heinegg, W. (2011). XVII Asymmetric Warfare: How to Respond?. International Law Studies, 87(1), 19.

Himelfarb, S. (25. Apr 2014). *Can Big Data Stop Wars Before They Happen? – Foreign Policy* . Načteno z Foreign Policy – the Global Magazine of News and Ideas : https://foreignpolicy.com/2014/04/25/can-big-data-stop-wars-before-they-happen/

Holford, Bill, ITProPortal. (01. Feb 2017). *Big Data on the battlefield – an introduction*. Načteno z ITProPortal | News & Guides For Business & Enterprise IT: https://www.itproportal.com/features/big-data-on-the-battlefield-an-introduction/

Holford, Bill, PublicTechnology. (08. Jun 2017). *How big data is helping to transform the defence sector*. Načteno z PublicTechnology.net | Digital, data and technology news, interviews and analysis for central and local government: https://www.publictechnology.net/articles/partner_article/bt/how-big-data-helping-transform-defence-sector

ICEWS. (2020). *ICEWS | Lockheed Martin*. Načteno z Lockheed Martin Corporation | Lockheed Martin : https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html

Idler, A., & Forest, J. (2015). Behavioral patterns among (violent) non-state actors: A study of complementary governance. Stability: international journal of security and development, 4(1).

IoBT REIGN. (2020). *Agile Synthesis (Composition) – Internet of Battlefield Things (IoBT) REIGN* . Načteno z Internet of Battlefield Things (IoBT) REIGN: https://iobt.illinois.edu/agile-synthesis-composition/

IoBT REIGN. (2020). *Internet of Battlefield Things (IoBT) REIGN* . Načteno z https://iobt.illinois.edu/

IoBT REIGN. (2020). *Internet of Battlefield Things (IoBT) REIGN* . Načteno z Demos – Internet of Battlefield Things (IoBT) REIGN : https://iobt.illinois.edu/demos/

Joshi, N. (26. Nov 2017). *Top 5 sources of big data | Artificial Intelligence | Data Science |* . Načteno z Application development | Big data| IoT | Digital Business | Cloud : https://www.allerin.com/blog/top-5-sources-of-big-data

Joshi, Naveen. (02. May 2017). *Big data Analytics: Descriptive vs. Predictive vs. Prescriptive*. Načteno z Application development | Big data| IoT | Digital Business | Cloud: https://www.allerin.com/blog/big-data-analytics-descriptive-vs-predictive-vs-prescriptive

Karlsrud, J. (2014). Peacekeeping 4.0: Harnessing the potential of big data, social media, and cyber technologies. In Cyberspace and international relations (pp. 141-160). Springer, Berlin, Heidelberg.

Laurence, M. (09. Jan 2020). *What Are the Benefits and Pitfalls of 'Data-Driven' Peacekeeping? – Centre for International Policy Studies*. Načteno z Centre for International Policy Studies – Canada's leading centre for analysis and debate of international affairs.: https://www.cips-cepi.ca/2020/01/09/what-are-the-benefits-and-pitfalls-of-data-driven-peacekeeping/

Letouzé, E., Meier, P., & Vinck, P. (2013). Big data for conflict prevention: New oil and old fires. *New Technology and the Prevention of Violence and Conflict*, 4-27.

Lynch, Clifford. Big data: How do your data grow? *Nature* [online]. 2008, **455**(7209), 28-29. ISSN 00280836. doi:10.1038/455028a

Manning, N. (08. Aug 2018). *Keeping the Peace - The UN Department of Field Service's and Peacekeeping Operations use of Ushahidi*. Načteno z Ushahidi: https://www.ushahidi.com/blog/2018/08/08/keeping-the-peace-the-un-department-of-field-services-and-peacekeeping-operations-use-of-ushahidi

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

Maoz, I., & McCauley, C. (2008). Threat, dehumanization, and support for retaliatory aggressive policies in asymmetric conflict. Journal of Conflict Resolution, 52(1), 93-116.

Mashey, J. R. (1998). *Big Data and the Next Wave of InfraStress Problems, Solutions, Opportunities | USENIX* . Načteno z USENIX | The Advanced Computing Systems Association : https://www.usenix.org/conference/1999-usenix-annual-technical-conference/big-data-and-next-wave-infrastress-problems

Mayer-Schönberger, Viktor. *Big data: a revolution that will transform how we live, work and think / Viktor Mayer-Schönberger and Kenneth Cukier*. 2013. ISBN 9781848547902.

Mazzei, Matthew a David Noble. Big Data and Strategy: Theoretical Foundations and New Opportunities. *Chapters* [online]. 2020.

mTrac. (2020). *Mission, Vision, Objective | Mtrac*. Načteno z Home | Mtrac : http://www.mtrac.ug/content/mission-vision-objective

Naeem, T. (21. Dec 2020). *Difference Between Structured, Unstructured & Semi-Structured Data* . Načteno z Enterprise-Ready Data Management Solutions | Astera Software : https://www.astera.com/type/blog/structured-semi-structured-and-unstructured-data/

Peace Inside. (2020). *Conflict prevention and early warning — Peace Insight* . Načteno z Peace Inside: https://www.peaceinsight.org/en/themes/conflict-prevention-early-warning/?location&theme=conflict-prevention-early-warning

PulseSatellite. (2020). *PulseSatellite: Human-AI Interaction for Satellite Analysis.* Načteno z UN Global Pulse - Big data for development and humanitarian action: https://www.unglobalpulse.org/microsite/pulsesatellite

Qatalog. (2020). *Qatalog • UN Global Pulse*. Načteno z UN Global Pulse - Big data for development and humanitarian action: https://www.unglobalpulse.org/microsite/qatalog

Ramsbotham, O., Miall, H., & Woodhouse, T. (2011). *Contemporary conflict resolution*. Polity.

Rao, V. R. (2018). *How data becomes knowledge, Part 1: From data to knowledge* . Načteno z IBM - United States : https://www.ibm.com/developerworks/library/ba-data-becomes-knowledge-1/index.html

Ricardo Real P. Sousa. The context of conflict resolution - international relations and the study of peace and conflict. *CEsA Working Papers* [online]. 2018.

Richards, Neil M. a Jonathan H. King. Big Data Ethics. *Wake Forest Law Review* [online]. 2014, **49**(2), 393-432. ISSN 0043003X.

Rossi, B. (03. Dec 2015). *Causation and correlation: a big (data) headache - Information Age* . Načteno z Information Age - For CTOs and Technology Leaders : https://www.information-age.com/causation-and-correlation-big-data-headache-123460611

SaaS Scout. (2020). *Big Data Statistics, Growth & Facts 2020 | SaaS Scout (formerly SoftwareFindr)* . Načteno z SaaS Scout - Product Research Platform powered by Big Data And AI: https://saasscout.com/statistics/big-data-statistics/

Sauer, T. (22. Feb 2017). *Take It from a European: NATO Is Obsolete | The National Interest*. Načteno z The National Interest: Defense, National Security, Politics and More: https://nationalinterest.org/feature/take-it-european-nato-obsolete-19537

Sean P. O'Brien. Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research. *International Studies Review* [online]. 2010, **12**(1), 87. ISSN 15219488.

Sector, I. T. S. (2012). Recommendation ITU-T Y. 2060: Overview of the Internet of things. Series Y: Global information infrastructure, internet protocol aspects and next-generation networks-Frameworks and functional architecture models. Retrieved from https://www. itu. int/rec/T-REC-Y, 2060-201206.

Seybolt, Taylor B., Humanitarian Military Intervention: The Conditions for Success and Failure (2007). Oxford: Oxford University Press, 2007, Available at SSRN: https://ssrn.com/abstract=1487382

Shachtman, N. (21. Jul 2011). *Exclusive: Inside Darpa's Secret Afghan Spy Machine | WIRED* . Načteno z WIRED: https://www.wired.com/2011/07/darpas-secret-spy-machine/

Simon, H. (1954). Spurious Correlation: A Causal Interpretation. Journal of the American Statistical Association, 49(267), 467-479. doi:10.2307/2281124

Southerton C. (2020) Datafication. In: Schintler L., McNeely C. (eds) Encyclopedia of Big Data. Springer, Cham. https://doi.org/10.1007/978-3-319-32001-4_332-1

Stevens, Marthe a Rik,Wehrens. Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data* [online]. 2018, **5** ISSN 20539517. doi:10.1177/2053951718816727

Swanström, N.L.P. and Weissmann, M. S., 2005, 'Conflict, Conflict Prevention and Conflict Management and Beyond: A Conceptual Exploration', Central Asia-Caucasus Institute and Silk Road Studies Programme, Washington, DC and Nacka, Sweden

The GDELT Project . (2020). Načteno z The GDELT Project : https://www.gdeltproject.org/#intro

UCDP. (2020). *Definitions - Department of Peace and Conflict Research - Uppsala University, Sweden* . Načteno z UCDP - Department of Peace and Conflict Research - Uppsala University, Sweden : https://www.pcr.uu.se/research/ucdp/definitions/#tocjump_13504946201092205_10

UCDP. (2020). *UCDP - Department of Peace and Conflict Research - Uppsala University, Sweden* . Načteno z Top ranked research and education - Uppsala University, Sweden : https://www.pcr.uu.se/research/ucdp/

UN Global Pulse, "Making Ugandan Community Radio Machine-Readable Using Speech Recognition Technology," Tool Series, no. 1, 2016.

UN Global Pulse. (2019). *UN Global Pulse Annual Report 2019 • UN Global Pulse* . Načteno z UN Global Pulse - Big data for development and humanitarian action: https://www.unglobalpulse.org/document/un-global-pulse-annual-report-2019/

UN Global Pulse. (2020). *UN Global Pulse - Big data for development and humanitarian action* . Načteno z About • UN Global Pulse : https://www.unglobalpulse.org/about/

UN Global Pulse. (2020). *Where We Work • UN Global Pulse* . Načteno z UN Global Pulse - Big data for development and humanitarian action : https://www.unglobalpulse.org/labs/

UN Peacekeeping. (2015). *Final Report of the Expert Panel on Technology and Innovation in UN Peacekeeping | United Nations Peacekeeping* . Načteno z United Nations Peacekeeping : https://peacekeeping.un.org/sites/default/files/performance-peacekeeping_expert-panel-on-technology-and-innovation_report_2015.pdf

UN Peacekeeping. (2020). *Terminology | United Nations Peacekeeping* . Načteno z United Nations Peacekeeping : https://peacekeeping.un.org/en/terminology

UN. (2001). *Extract from the Report on the World Social Situation, 2001.* Načteno z DISD | The Division for Inclusive Social Development (DISD) is part of the Department of Economic and Social Affairs (DESA) of the United Nations Secretariat.: https://undocs.org/E/2001/104

Van Der Sloot, Bart a Sascha Van Schendel. Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study. Journal of Intellectual Property, Information Technology and Electronic Commerce Law (Bridle, 2018)online]. 2016, 7(2), 110-145. ISSN 21903387.

Wallensteen, Peter.Understanding conflict resolution / Peter Wallensteen. 2012. ISBN 9780857020499.

Ward, Jonathan Stuart a Adam Barker. *Undefined By Data: A Survey of Big Data Definitions*[online]. 2013.

Ward, M. D., Beger, A., Cutler, J., Dickenson, M., Dorff, C., & Radford, B. (2013). Comparing GDELT and ICEWS event data. Analysis, 21(1), 267-97.

Weinberger, S. (2017). *The imagineers of war: the untold history of DARPA, the Pentagon agency that changed the world*. Knopf.

Whaley, R. (09. Aug 2019). *The big data battlefield - Military Embedded Systems* . Načteno z Home - Military Embedded Systems : https://militaryembedded.com/ai/big-data/the-big-data-battlefield

WHO. (09. Jan 2014). *Global status report on violence prevention 2014*. Načteno z WHO | World Health Organization: https://www.who.int/publications/i/item/9789241564793

Yonamine, J. E. (2013). Predicting future levels of violence in Afghanistan districts using GDELT. Unpublished manuscript.

Zwijnenburg, W., & Waleij, A. (13. Jan 2016). *Fire and Oil: The Collateral Environmental Damage of Airstrikes on ISIS Oil Facilities*. Načteno z New Security Beat - The blog of the Wilson Center's Environmental Change and Security Program, with contributions from the Maternal Health Initiative, China Environment Forum, and more:

https://www.newsecuritybeat.org/2016/01/fire-oil-collateral-damage-airstrikes-isis-oil-facilities/