Big Data in ISR

Big Opportunity for Data Analysis Challenges

By Major Giuseppe Valentino, IT AF, JAPCC

Introduction

Over the past few decades, the use of data and information has been exponentially increasing. While data collection is an expected and thus normalized facet of life, reinforced by the many bureaucratic protocols people adhere to in order to live, work and travel, it is profoundly different in military operations, where data processing is the key to achieving effective results. The overwhelming volume of data, combined with its complexity, causes any necessary decision-making to be a long, drawn out and cyclical process which impacts ongoing operations. Indeed, the human brain lacks the capacity to manage information in a short-time frame and find the appropriate response quickly.

For these reasons, the Big Data issue, related to new technologies and modern multi-domain operations, has been a critical focus area in recent NATO debates. Current theories are unable to define the concept of Big Data coherently or consistently, mainly due to the complexity and ambiguity of the idea[1]. The aim of this article is to cast light on Big Data and its potential purposes in the Intelligence Surveillance Reconnaissance (ISR) in terms of the data management approach.

This article reflects a portion of a wider JAPCC study on the same subject. Readers are encouraged to look at the 'Big Data Management in ISR and New Technology Trends' white paper for a more in-depth understanding.

Big Data Statement

The term Big Data[2] can be defined as multiple sets composed of many bits of data, analyzed or not, and interrelated by tools based on dedicated algorithms for information exploitation. It might not involve a new technology or a new database, but 'it is a relationship between data and the organizational, procedural and cultural factors that make up the enterprise. It is primarily about the ways in which data is managed.'[3] According to the Gartner definition, 'Big Data is 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.'[4] The statement quoted in this article most represents the current view in supporting this explanation.

Big Data is comprised of collected datasets and information that are characterized by complexity and are not immediately useful for exploitation. The simple and disaggregated dataset may not support the decision-making process at any level. Overall, the use of Big Data could bring an operational advantage, but only if the information can be turned into readable data in a time-sensitive manner to deliver indications and warnings for threat analysis.

The scientific world defines Big Data by five leading attributes:[5]

Volume: Currently available data reaches into the terabytes and even petabytes. The 'big' represents millions and billions of information bits stored in different databases. To quantify the exact volume of data is difficult to do at any one moment.

Velocity: Refers to the data exchanged among the interconnected systems within a specific measure of time. In daily life, organizations constantly and continuously collect new data, in real-time, from myriad sources, and exchange information among organizations very rapidly.

Variety: Big Data is comprised of text, images, videos, and other information continuously created, collected, and shared.

Veracity: Accuracy is not always present in large data or datasets. Acquiring a huge variety of data can reduce the level of accuracy. Accuracy is a criterion to select and clean data to determine what portions should be processed to obtain useful information.

Value: The term 'value' is often used interchangeably and without precision. It might represent additional information that it is achievable only by combining a huge amount of data by tools[6]. Big Data is as valuable as its utilization in the generation of actionable information. There is no disagreement that data holds the key to actionable insights to validated information; however, the post-modern organization needs to progress quickly to be able to analyze data automatically. Moreover, it is essential to understand the patterns within the data, and to provide solutions in visual[7] and readable information to add value.

Overall, Big Data could be considered a federated database[8] with the following features:

• Structured data;
• Unstructured data;
• Semi-structured data.

The challenge is to enable automatic data processing and data interconnection among information streams for advanced applications. The correct use of Big Data, through employing different information or databases sourced from all domains, may be essential to reach the desired information dominance to support military operations.

The 'world of data' provides the possibility of having a copious amount of differing information simultaneously. The current dilemma is how to obtain usable intelligence to support Commander's Critical Information Requirements (CCIR)[9] and related priorities inside the complexity of the decision-making process. Recently, many nations have improved their capability to gather information due to the possibility of interconnecting multiple databases. Despite the increasing capabilities of available systems, a critical point remains as to the ways in which to manage the countless amounts of data acquired. In other words, complex organizations, like NATO for example, should

© geralt/Pixabay.com

be able to establish a common and standard criterion to use data and related analysis. It may be reasonable to follow the systematic approach by providing for common:

• Data strategies;
• Management of data systems;
• Policies for data storage;
• Validating processes and algorithms.

Management of priorities became the first element inside the Big Data analysis that was essential in supporting operational data users.

Another important priority is the data storage, considering that much of the information about a potential adversary is already present in the databases. Nevertheless, information should be stored according to a useful structure, which guarantees and implements basic intelligence practices. Furthermore, the most critical aspect for NATO is to have sharable ISR databases that could support information systems and warnings.[10] The methodology and data matching should be a pre-defined process to validate the analysis.

Using Big Data

Recently, NATO acquired its first owned unmanned ISR system[11], Global Hawk, which will permit the gathering of videos and data worldwide. In fact, ISR can count on support from Air Power and Space through images and other data collection that is essential in obtaining current intelligence as well as maintaining information superiority. Moreover, the NATO information system is integrated by other data sources and national data contributions based on a federated architecture. These outstanding volumes of data highlight the opportunity to identify a new architecture of technology to enhance the use of Big Data to optimize the Alliance's capabilities in intelligence management.

Recently, researchers displayed an increased interest in new technologies for Big Data, which provides an important opportunity to apply and optimize data fusion process for military purposes. At the same time, it is reasonable to note the data analysis in which the key factors are there to protect allied assets and troops. The use of emerging technologies and tools represents a turning point in transforming intelligence analysis and command and control synchronization.

As argued by Air Chief Marshal, Sir Stuart Peach, '[Big Data] holds great potential for the Defence and Security sector but [the Ministry of Defence] must not fall into the trap of procuring only bespoke software solutions if it is to exploit the technology in a timely manner.'[12] The challenge in ISR is to apply Big Data with a clear management methodology for reducing the time taken to disseminate information and mitigate errors in assessments due to human factors.

Improving the Processing Exploitation Dissemination Cycle to Support Commander's Decisions

ISR issues in NATO are well known and have been highlighted in recent operations. The limitations due to data exploitation in ISR manifested during the crisis in Ukraine where NATO was surprised by a Russian so-called 'snap exercise'. Information was collected, but not immediately available, for actionable intelligence to support rapid decision-making on how to exploit the information gathered on preventing Russia actions in Eastern Ukraine. The key to success for the future of ISR is to synchronize operations and intelligence by continuous exploitation of analyzed data, ensuring it is reliable and validated. In the Processing, Exploitation and Dissemination (PED) cycle, ISR should guarantee the full integration of operational domains and set up a new mindset based on multi-domain data collection. In other words, Big Data brings a new perspective inside PED, in terms of timely, validated and actionable 'readable data' that need to be defined.

Currently, NATO counts on multiple databases and various information datasets, managed by various nations; the critical link in this system is the missing interconnection of those databases between storage locations as well as a lack of optimized data processing in the overall exploitation process. The old concept of PED[13] should be revisited from a new perspective, where the time and reliability of gathered information and its reliability to guarantee a seamless transition and translation into the decision-making cycle is assessed. One of the important things to consider is the role of networking among systems based on cloud computing resources, capable of collecting data and processing information close to the relevant 'event'[14].

Task Collect Process Exploit and Disseminate (TCPED) at the Tactical Level

At an operational and strategic level, data could be analyzed, correlated with other sources and evaluated accurately, while the tactical level manages the huge amount of incoming information, which would translate to data being made available in a timely manner in current operations to support real-time and near-real time decision making. Within the current intelligence management structure, in which the key to success is the time taken to gain awareness of the ongoing situation, Big Data's strategy plays an essential role in generating actionable intelligence. The high response time required to detect and collect data before processing and dissemination, is a real challenge that ISR needs to face. From this perspective, Big Data in the ISR environment will facilitate an understanding of what is valuable from the basic intelligence and provide information to extrapolate these findings into the current situation. This ensures exploitation of data reliably and in an automatic manner.

To illustrate this point, the use of tools for Moving Target Identification may have been imagined, in which a pre-defined set of data could match a variety of target information and other related sources referenced to obtain usable intelligence and deliver assessments in the shortest time frame possible. The 'Unified Vision' exercises, based on federated data exploitation architecture, underlined the notion that TCPED[15] is a critical issue due to the lack of alignment and standardization between datasets. It is presumed that in information analysis, 88 % of 'usable data is being left untouched'[16] due to a multitude of characteristics such as data complexity, human capacity, storage issues and connectivity. There is no doubt that emerging technologies (e.g. Machine Learning, AI, etc.) will continue to progress in the future together with advanced tools for optimizing intelligence exploitation, but Big Data strategy is the first step in
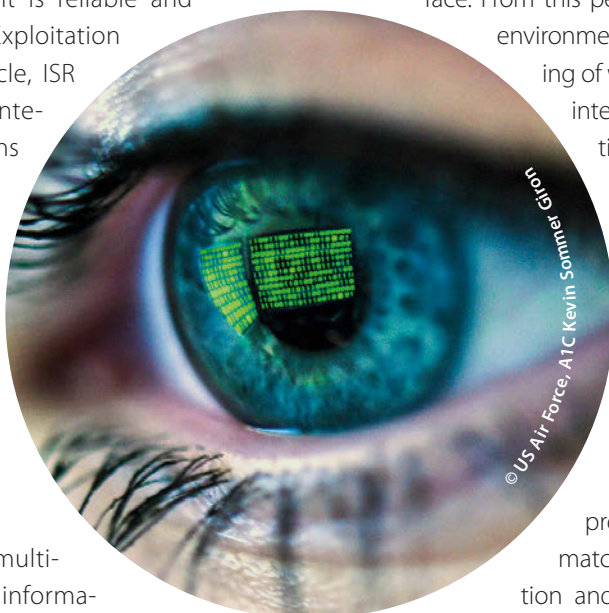
© US Air Force, A1C Kevin Sommer Giron

sustaining intelligence applications and resolving the information dilemma for the decision-making process.

Recommendations

In conclusion, the following recommendations should be considered in order to apply the Big Data concept and put in place the required methodologies to orientate it to military use and exploitation:

- Applying standard concepts for linking databases from the Intelligence disciplines;
- Interconnecting structured and unstructured databases by software applications and algorithms;
- Selecting information more easily, quickly and accurately;
- Structuring tools and queries according to Intelligence Requirements Management (IRM);
- Data matching and fusion by analysis;
- Enhancing information from text, documents, raw data, crypto information and converting them into actionable intelligence;
- Extracting data, building models and delivering intelligence solutions;
- Avoiding arbitrary systems of data classification;
- Reducing the complexity of information;
- Optimizing multiple-source data fusion.

It is important to define the correct algorithm based on a clear Big Data strategy for ISR purposes. Big Data is synonymous with 'big opportunity'. The traditional ISR PED system has proven itself to become overwhelmed by the sheer volume of data and it is necessary to build alternative approaches in terms of analysis, data storage and information sharing. The first step is to identify NATO's Big Data strategy through a multi-disciplinary and multi-domain vision. The overall aim is to reduce data complexity and consistently support the decision-making process at all levels of the command structure. ●

1. NATO Science&Technology Organization (2020), Science and Technology trends – Exploring the S&T Edge, p. 13. Available from: https://www.nato.int/nato_static_fl2014/assets/pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf (accessed 22 Feb. 2021).
2. Col Shane P. Hamilton, Lt Col Michael P. Kreuzer, (2018), Big Data Impreative. Available from: https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Volume-32_Issue-1/F-Hamilton_Kreuzer.pdf (accessed 20 Jan. 2021) The characteristics of Big Data inside military environment has not defined yet. For instance, The Big Data in the article 'Imperative Air Force Intelligence for the Information Age' written by Col Shane P. Hamilton, USAF Lt Col Michael P. Kreuzer, USAF, the Big Data have been analysed according to the USAF standpoint and features in 4 characteristics.
3. Peter Roberts and Andrew Payne (2016), Intelligence, Surveillance and Reconnaissance in 2035 and Beyond. Available from https://rusi.org/sites/default/files/201602_op_isr_in_2035_and_beyond.pdf (accessed 22 Feb. 2021).
4. Gartner, Inc., https://www.gartner.com/en/information-technology/glossary/big-data (accessed 22 Feb. 2021)
5. Bernard Marr (2019), Why only one of the 5 Vs of big data really matters. 19 March 2019. Available from https://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters (accessed Dec. 2020)
6. Laurenz Wuttke (undated), What are Big Data, 4 V's and technology, available from: https://datasolut.com/was-ist-big-data/#big-data-4-v (accessed 8 Mar. 2021).
7. Bryan Harris, 'Closing the OODA Loop: Using Big Data and Analytics to Improve Decision Making Insights From a Military Operations Research Society (MORS) Industry Showcase', SAS, 2015 p. 7. 'Data visualization is the presentation of data in a pictorial or graphical format. Even when data volumes are very large, visualization allows people to spot patterns that were not obvious to them before quickly and easily. Visualizations convey information in a universal manner and make it simple to share ideas with others. Yet Harris pointed out that visualizations only go so far in communicating results: 'How many people have seen charts, nodded and pretended they understood them?' Visualizations and the dashboards that present them are important, but insufficient. Explained Harris, 'Visualizations are evidence. But at the end of the day, people remember stories, not data'. Available: https://dsimg.ubm-us.net/envelope/364533/392343/ClosingTheOODALoop.pdf.
8. Bianca Barbu, David Martin, Lora Hadzhidimova (2016), Advances in Defence Analysis, Concept Development and Experimentation: Innovation for the future. Available from https://www.act.nato.int/images/stories/media/capdev/capdev_03.pdf (accessed 21 Feb. 2021).
9. AJP 2.7, Allied Joint Doctrine for Joint Intelligence, Counter-Intelligence and Security (NU), Edition A Version 2, Feb. 2016, p. 5-1.
10. Ibid.
11. PAO AirComm (2020), NATO Alliance Ground Surveillance Controls Firs RQ-4D Flight. Available from https://ac.nato.int/archive/2020/1st_NAGSF_Tng_Flight (accessed 22 Feb. 2021).
12. Neil Couch and Bill Robins (2013), Big Data for Defence and Security. Available from https://rusi.org/event/rusi-occasional-paper-launch-event-big-data-defence-and-security (20 Jan. 2021).
13. AJP 2.7, Allied Joint Doctrine for Joint Intelligence, Surveillance and Reconnaissance (NU), Edition A Version 1, Jul. 2016, p. 1–3.
14. Frank T. Johnsen, Norwegian Defence Research Establishment (2017), STO-MP-IST178 'Towards Big Data in the tactical domain'. Available from https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-178/MP-IST-178-11.pdf (accessed 20 Feb. 2021).
15. Ibid. 9, p. 3-1.
16. Altair, Connecting the dots: how data prep addresses three major challenges for the Intelligence community to improve threat detection, Sep. 2019, p. 3. Forrester (2014), Big Data Hadoop Solutions. The Forrester Wave.

Major Giuseppe Valentino

started his career in 1992 joining the Italian Air Force NCO School in Caserta. During his service worked in E.W. group as an Intelligence Analyst and had two tours to Sarajevo (Bosnia-Herzegovina) and was responsible for analysis and production in the Deployed Intelligence Cell.

In 2005, he was engaged as a Force Protection platoon leader in Kosovo as well as supporting other key NATO operations. From 2010 to 2019 he was Section Head of COSMO-SkyMed operations in Italian Defence User Ground Segment (IDUGS). Major Valentino holds an MA with honours, in Political Science Sapienza University of Rome and an MA (level II) in Peacekeeping and Security Studies University Roma Tre, and currently serves as an ISR Subject Matter Expert at the JAPCC.