Joint Analysis and Lessons Learned Centre

NATO's Lead Agent for Lesson Learned

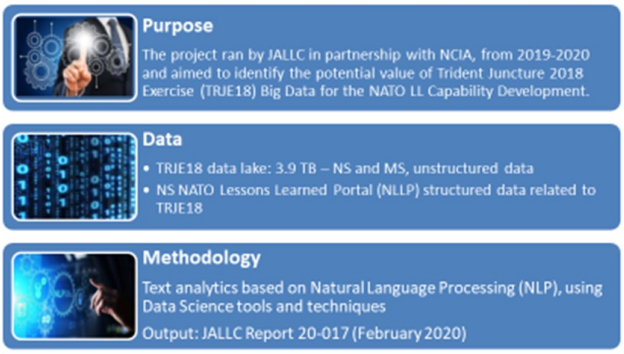
JALLC PROJECT FACTSHEET

NATO Exercise Big Data Exploration

Report Published 13 February 2020

Project Overview

In an increasingly digital world, NATO's reliance on data, technology and connectivity is growing and generating a need for constant innovation. In this dynamic environment, it is vital for the Alliance to detect, record, learn, and share lessons at the speed of relevance. However, studies have shown that, particularly within NATO Exercises, relevant Lessons Learned (LL) information is often lost in *Big Data* - large volumes of diverse data generated at a fast pace.



Overview of the NATO Exercise Big Data Exploration Project

> **"The results from this project highlight the fact that the nature of Lessons Learned data requires human input to maximize the value of data for positive outputs in the field"**

Data Science tools and techniques can help address some of the challenges posed by Big Data. For NATO exercises, these tools can revolutionize the way we identify and extract relevant information, making it more efficient and cost-effective. To test these ideas, the JALLC embarked on a project in 2019 to explore the data collected at Exercise NATO Trident Juncture 2018, the biggest Live Exercise NATO has conducted since the end of the Cold War. In partnership with the NCIA Data Science Team, the JALLC investigated how Big Data Analytics could help the NATO LL Capability development by: processing available data faster and more effectively, accessing necessary information more efficiently, and sharing relevant information more easily.

The Data Science Process

The report provides more detail on the results of this exploration, models of investigation, the Data Science techniques applied, the challenges faced, and the solutions obtained.

Promising Results and Challenges

The project results demonstrated that leveraging Data Science tools and techniques can help improve the implementation of the NATO LL Capability. The techniques used were helpful in identifying and extracting Lessons-related information from the exercise data by looking for keywords and specific entities. One type of technique, sentiment analysis, classified existing language as positive, negative or neutral in order to identify information pointing to a problem (a potential *Lesson Identified*) or a success (a potential *Best Practice*) discussed during the exercise. While promising, the results need to be further refined to improve their overall accuracy and relevance.

Sentiment analysis overview dashboard

Utilizing these tools could result in immediate short term benefits, including faster software data processing and improved information visualization via an interactive and customizable dashboard. The dashboards make the data more accessible, while also formatting and standardizing the existing items in the NATO Lessons Learned Portal.

In the long term, predictive modelling could unlock more hidden value in the data. The project results showed that machine learning algorithms could be used to predict observation categories based on a content text analysis. This could potentially address the issue of missing or incorrect metadata for lessons already in the Portal, improving the Portal's search functionality, and opening the door for more user-friendly applications.

Ultimately, NATO can greatly benefit from these Data Science tools, which are capable of processing large amounts of data coming from operations and exercises, in a short timeframe. However, to leverage these technologies, a number of challenges must be addressed, ranging from barriers to data acquisition and use, to the absence of specialized tools that are able to transform highly technical NATO data into computer readable formats. The nature of LL data also requires human input to maximize the value of data for positive outputs in the field. While machines can help automate key parts of the data collection process, extraction, and analytics, the human aspect remains central to performing the more complex analysis behind the lessons.

Initial Recommendations

To harness the Data Science tools and techniques for LL, the project team made a series of recommendations designed to keep the momentum for innovation. The initial recommendations aim to ensure that the preliminary conditions are in place to realize the value of Big Data: removing barriers to data collection, training JALLC staff in using Data Science tools and concepts, and creating a specialized tools for LL to improve the quality of results.

Project Team

Dr. Mihaela Racovita, JALLC
ROU NIC, Project Manager

Dr. Mihaela Racovita is a Research Analyst, and holds a PhD in International Relations from the Graduate Institute of International and Development studies in Geneva. She joined JALLC in 2019, managing the NATO Exercise Big Data Exploration Project, and supporting work on COVID-19 lessons. Prior to joining NATO, Dr. Racovita worked as a researcher on small arms control, gender, and armed violence, and collaborated with the UN, African Union, ECOWAS and the OSCE.

Ms. Jacqueline Eaton, JALLC
GBR NIC, PORA

Dr. Michael Street, NCIA
GBR NIC, Head Innovation and Data Science

Ms. Ivana Ilic Mestric, NCIA
HRV NIC, Senior Data Scientist

Mr. Arvid Kok, NCIA
NLD NIC, Senior Data Scientist

Mr. Giavid Valiyev, NCIA
ITA NIC, Junior Data Scientist

Ms. Jodie Lazell, JALLC
GBR NIC, Editor

If you are interested in this or any other JALLC Analysis product, please contact the JALLC.

JALLC

Phone: +351 21 771 7007/8/9
Fax: +351 21 771 7098
E-mail:jallc@jallc.nato.int
www.jallc.nato.int
Visit the Portal: https://nllp.jallc.nato.int

Avenida Tenente Martins
1500-589 Lisbon
Portugal

A proud member of Allied Command Transformation