Large Scale data processing & open source intelligence

Context

COM JFC's Future View Paper: *Warfare in the Information Age* sets out JFC's view on the key trends which will impact on Defence operations in the future and what measures JFC needs to put in place in order to face future threats. A key theme within this paper is that: *".... information-centric capability employed in information-centric operations can ameliorate many of the shortcomings of a reducing number of platforms and people."*

In order to achieve this goal MOD requires advice and evidence to understand the most effective information and intelligence (I2) processing capabilities to meet needs across all Defence users in the future.

Recent operations have demonstrated the operational value which MOD can gain by integrating open source information and intelligence into a wider UNDERSTAND[1] capability, with significant investment now being placed by Joint Force Command in developing a future OSINT system. MOD requires support from Dstl in understanding both the processing capabilities and the data sources which can provide greatest value, delivered through applied experimentation and demonstration.

DCDC Global Strategic Trends (5th Edition) outlines a future where technology will continue to rapidly develop, with an increased number of connected devices globally. This is likely to have a knock on effect to MOD operations with an increase in I2 sources in terms of volume, variety and velocity. To effectively manage this increase in available intelligence, MOD requires support in understanding the value and DLOD implications of emerging technologies which are designed to handle large scale data processing (e.g. The Apache Hadoop Framework).

Project Aim

Applied research delivered as a time-bounded project (2 years) to help MOD to understand the potential for emerging information technology to transform the way we provide military capability and conduct operations in the future; as well as research to inform future pan-Defence OSINT capability delivery, as part of the Equipment Programme.

Project Objectives

- The project will deliver the S&T and Innovation in support of MOD's future OSINT enterprise, in order to ensure that future capabilities meet all user requirements, are value for money, and integrate cutting edge technologies;
- The project will deliver the evidence and understanding of the benefits of Big Data technologies to MOD problems; how these can be integrated as part of a wider ISR architecture; and the associated DLOD implications, so that stakeholders can make an informed decision on future equipment procurement.

OSINT

Over the next 2 years MOD will see a significant enhancement to its current OSINT capability.. This WP will be the single focal point for research in support of MOD's future OSINT capability, with direction being provided by the OSINT CIWG, which should be regularly attended by the research team. The principle aim of the WP is to ensure that future OSINT capabilities delivered to MOD meet all user requirements including JFIG, DIAS, and Front Line Commands (FLCs). This should be achieved through an applied research activity which looks to understand the wide breadth of application areas and benefits of OSINT within Defence.

Large Scale Data Processing

MOD struggles to handle the information it currently collects in an effective manner; a situation which is likely to worsen in the future due to the increased volumes of data being generated, and the wide range of potential sources in which they may come from. MOD needs to understand the benefits which emerging "Big Data" technologies can provide, and how

[1] UNDERSTAND is part of MOD Doctrine

they are best applied to Defence problem-sets, in order to support decision makers. Correct implementation of such technologies will enable MOD to work towards a scalable and robust intelligence processing capability which is resilient to increases in the volume, variety and velocity of data sources. This WP will address a broad set of requirements across MOD to help users understand the benefits; how these technologies could be integrated as part of a wider ISR architecture; and the associated DLOD implications.

In the context of this document the following should apply when deciding what is and isn't within the scope of the WP:

- If the data (in its entirety) can be processed on a single workstation/laptop within operational timescales then it is not large scale;
- If traditional RDBMS based technologies are better placed to process the data then it is not large scale.

There is currently no exploitation path for a "Big Data" capability for MOD and therefore this WP will have a greater flexibility in scope and research direction. Focus should be on experimentation and demonstration of state of the art in order to understand the possible benefits for Defence. Given the wider interest and potential benefits offered by Big data technologies across MODs business this WP should not be constrained by purely focussing Intelligence use cases.