"Intelligence in a Data-driven Age by Cortney Weinbaum and John N.T. Shanahan"

Ranjan Prabhu

(*Cortney Weinbaum and John N.T. Shanahan, "Intelligence in a Data-Driven Age," Joint Forces Quarterly 90, 3rd Quarter 2018, pp. 4-9.)*

Accessing and analysing news from around the world has become the key to understanding the global security environment. In view of this, in 1941, United States (U.S.) President Franklin D. Roosevelt established an office called the Foreign Broadcast Information Service (FBIS), to be run out of the Central Intelligence Agency (CIA). The office's mandate was simple: translate the news from around the world for US policy-makers to make informed decisions. However, in the past, there was only a handful of electronic and print media which needed to be scoured for news by analysts. But, in the current times, the situation has changed drastically. In today's world, there is a multitude of news platforms which produce an immense amount of news data that require a large number of analysts to process the information—making such an organisation uneconomical and unwieldy.

Colonel **Ranjan Prabhu** is an instructor at the Faculty of Gunnery, School of Artillery, Devlali.

Given this perspective, the evolving challenge is not only to identify the news sources which can provide the required data but also to predict future outcomes, based on millions of bytes of data which are continuously generated over a 24-hour cycle. In a world where every human is interconnected with the world 24x7 and is emitting data through social media platforms, browsing habits, blogs, vlogs and other such media, data is produced at a very high rate. This then requires a huge amount of processing power to sift through and analyse the available data into valuable information. In this regard, data science integrated with Artificial Intelligence (AI) has significantly succeeded in sifting through this mass of multi-source, multi-language , disparate and scattered datasets while accurately predicting events which might occur based on the 'tones' of these emitted data on social media and websites.

The Global Database of Events, Language, and Tone (GDELT), created by Kalev Leetaru of *Yahoo!* and Georgetown University, along with Philip Schrodt and others, describes itself as:

> An initiative to construct a catalogue of human societal-scale behaviour and beliefs across all countries of the world, connecting every person, organization, location, count, theme, news source, and event across the planet into a single massive network that captures what's happening around the world, what its context is and who's involved, and how the world is feeling about it, every single day.[1]

This database used millions of publicly available records, robotically analysed for tone along 1,500 dimensions, to accurately pinpoint the location of Osama bin Laden within a 200-km radius of where Bin Laden was eventually found in Pakistan. The co-founder, Kalev Leetaru, used the datasets provided by the database to effectively predict the events of the social or civil uprisings such as the Arab Spring.[2] With the availability of such powerful tools in the public domain, it is not surprising that there

is huge interest and clamour amongst the Intelligence Community (IC) in data analytics and data science, in order to be able to predict events which may be of national security concern.

In this backdrop, the article, aptly titled "Intelligence in a Data Driven Age" by Cortney Weinbaum, a management scientist at the RAND Corporation and Lieutenant General John N.T. "Jack" Shanahan, US Air Force (US AF), Director for Defence Intelligence (Warfighter Support) in the Office of the Under-Secretary of Defence, published in the *Joint Forces Quarterly*, a publication of the National Defence University of the US, critically examines the importance of data emanating today from a multitude of disparate sources and, how it can drive the IC to correctly predict events of concern to national security.

Asserting that the vector, volume, velocity, variety and ubiquity of data are disrupting traditional tools and methods of national security policy, operations, and intelligence, the authors state that the failure to treat data as a strategic asset will cede precious time and space to competitors and adversaries. Spelling out the peculiarities of IC data,[3] Weinbaum and Shanahan argue that such data can be organised, sifted and analysed in critical timeframes only by embracing machine learning tools and devising policies, which encourage human machine teaming to flourish. In this respect, describing the future battle space as consisting of algorithms, networks and sensor grids in addition to ships, aircraft, tanks and missiles, the authors pointedly note that future wars would be fought on civilian and military infrastructures of satellite systems, communication and electrical grids, transportation systems and human networks. Therefore, monitoring of all such critical infrastructure would then necessitate analysing data of disparate data sets, which is a task most suited to AI based analysing tools.

Noting that U.S. does not have a broad national strategy on AI, Weinbaum and Shanahan reflect on the key concern, highlighting that China is likely to outmatch the US by 2030 in this field of utilisation of AI for data analytics. This concern spelled out by senior researchers in the US

IC and think-tanks is a matter of immediate attention for policy-makers in India, which has not so cordial relationship with respect to border issues with China. As the authors correctly assess, though data is increasingly being collected through a larger number of multiple types of sensors than ever before, unless a clear policy path on how such data need to be analysed is drafted, it will fail to provide any valuable information to the IC. To which, creation of huge data warehouses with aggregated data is a better path for obtaining intelligence of value in today's technological and data driven scenario.

While acknowledging that intelligence of value can be obtained using modern data analytic tools, the authors contend that such information can be useful only if appropriate combat systems are developed, which can then utilise it to decide a course of action and act accordingly. To which, Weinbaum and Shanahan in their assessment, conclude that AI based systems would be the right choice in the near future, ensuring that each step of the Observe Orient Decide Act (OODA) loop is accelerated in a manner that can outsmart the adversary. In this regard, with most nations developing such systems, future wars may well be algorithm versus algorithm wars wherein decision-making may parse into milliseconds—a capability which is beyond the realms of human ability.

It is known that AI-based intelligence analytics can fuse and analyse data at rates faster than any human can; the fact that all types of data then need to be accessible to such systems is a key issue that the IC needs to deal with. In an environment where data is treated with total confidentiality, with accessibility being highly controlled, allowing an algorithm access to all data then opens up the potential risk of manipulation by adversaries. This risk, therefore, requires intelligence analysts to be trained in recognising attempts by adversaries to manipulate or alter data—which has been correctly assessed by the authors.

Weinbaum and Shanahan in their study strongly recommend that any AI-based analytics programme or system made for the IC must be

based on open architecture in order to be adaptable to fast changing technologies and to be able to ingest multi-source, multi-type data. They also support the fact that in today's world, much of the data which may be relevant to an IC, would be coming from open sources and, therefore, it deems to challenge the very concept of 'intelligence' wherein, classification of information is a parameter to judge its value. This truism was demonstrated during the Crimea crisis in 2014, when much of the information about Russian participation in the annexation was actually derived from the data trail mainly left behind by publicly sourced and available information.[4]

The vast scope of analytics, therefore, needs a skilled workforce, adept at understanding activity patterns rather than individual pieces of information. As Weinbaum and Shanahan rightly point out, adaptation of AI analytics has to be supported by revolutionary changes in human capital, technology acquisition processes and Research and Development (R&D). While technological change in the industrial age occurred at a moderate tempo and as a result, military doctrine was based on the fundamental premise of mass production, future warfare will feature a myriad technological advances that come at a tempo that disallows mass production. Therefore, the authors conclude that the system needed to sustain AI-based intelligence analytics should be based on the "prototype warfare" concept, which caters for quick absorption of niche technologies at much faster rates, thus, creating an asymmetrical advantage against an adversary. While fully supporting the need to infuse high-end technology into intelligence analysis, the authors significantly note that a healthy mix of traditional human intelligence gathering tradecraft needs to be maintained within the US IC. This will then help it to align with the requirement while fighting a low tech adversary, like the one that the U.S. is faced with in Afghanistan.

Weinbaum and Shanahan have emphasised on the need of infusing machine learning in the IC. More specifically, in the Indian context, this

imperative has become the need of the hour. For India, at one end, is faced with a belligerent adversary on the western front, which uses terrorism as a tool of its state policy. While at the other, it deals with a far more assertive neighbour on its northern borders, whose closed society provides very little intelligence. Given these challenges, data from multiple sources which constantly emanates from an interconnected world, can only be amalgamated to study patterns at rapid rates if AI and machine learning are adopted. It is only then that the prediction of future events having strong national security implications can be carried out accurately and in a timely manner which will help us stay within the adversaries' OODA loop. In an overall assessment, the article is futuristic in its concept and recommendations and is a must read for India's intelligence community, policy-makers and students of military affairs.

Notes

1. See Wikipedia, "Global Database of Events, Language, and Tone," https://en.wikipedia.org/wiki/Global_Database_of_Events,_Language,_and_Tone#cite_note-1. Accessed on March 29, 2019.
2. Patrick Tucker, "How the Internet Could Have Predicted the Invasion of Ukraine," *Defence One,* April 14, 2014, https://www.defenseone.com/technology/2014/04/how-internet-could-have-predicted-invasion-ukraine/82480/. Accessed on March 29, 2019.
3. Pecularities comprise data that is generated in too many diverse formats, in too many disconnected or inaccessible systems, without standardised structures and without overarching agreed-upon ontology.
4. Jason M. Brown, "The Data Driven Transformation of Intelligence," *The National Interest,* February 25, 2017, https://nationalinterest.org/blog/the-buzz/the-data-driven-transformation-intelligence-19570. Accessed on March 30, 2019.