University of Exeter

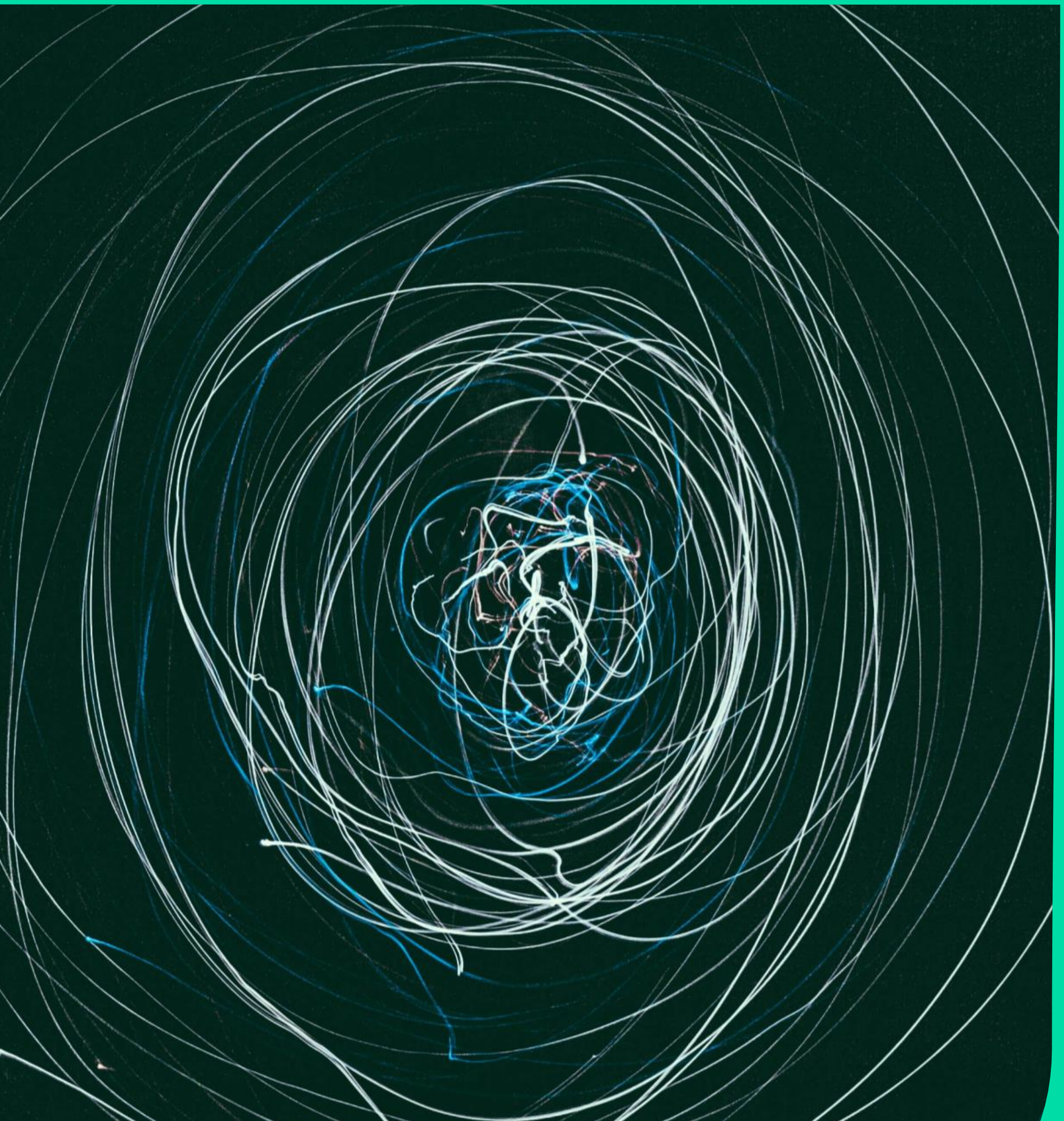
Policy @ Exeter

RESEARCH NETWORKS

Invisible Combat

The journey into the emergence of machine intelligent processes in contemporary conflict

Elizabeth Vallance-Bull



Policy @ Exeter

Invisible Combat
Elizabeth Vallance-Bull

This Centre of Public Understanding of Defence and Security briefing paper is a lightly edited version of a presentation made by Libby Vallance-Bull to the House of Lords Committee on 1 May 2024. Libby is Hon Senior Research Fellow at CPUDS.

This story starts in 2002, not in defence and national security, but in oil and gas exploration. A seismic dataset belonging to BP, terabytes in size from offshore West Africa, had just been delivered to the office of a company where I was employed as a geophysicist.

The aim of our work was to find a target – a reservoir, oil bearing sands - and establish their extent, then to drive 'oil in place' calculations, and ensure the economic value of the hydrocarbons for which we would be spending hundreds of millions on drilling. Drawing on the same techniques used to detect cancerous cell anomalies in human MRI scans, we set about our pioneering work. We were creating and refining high-definition frequency decomposition algorithms, fused with red, green and blue (RGB) blending, and iso-proportional slicing to create meaningful structural and stratigraphic multi-attributes. I use the technical language neither to confuse nor impress, and will return to this point later; it is important as expert practitioners that we use the right language from the get-go.

I remember very clearly the moment we revealed the results or our work (the hours of model refinement and iterations on our calculations) We visualised the target on a screen in high fidelity and granularity; the hydrocarbon bearing reservoir, its extent, as well as its geospatial and temporal context and relevance to the rest of the world.

We had used machine intelligent processes, on terabytes of data, to find a target deep in the subsurface, hidden to the human eye, using high performance computing power.

Fast forward 19 years to 2021, Year 2 of COVID. James, a colleague, is in the back of a London Black Cab, at that time working on the UK COVID vaccination scheme (UKCVS). I am in a secure facility in London, and we are talking online with a company who were demonstrating software which identified mis-, dis- and mal- information from very many social media outlets across the globe.

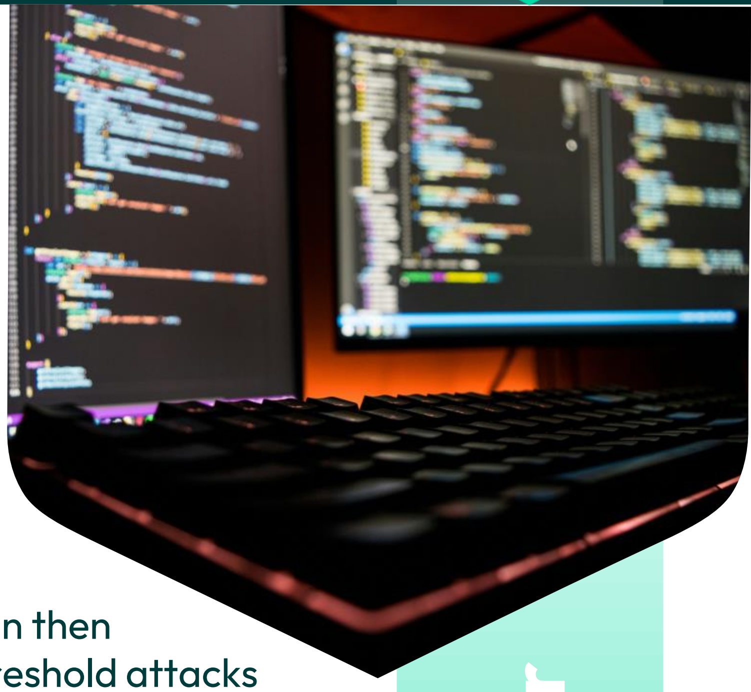
The aim of our work on this occasion was to find anomalies in millions of open-source intelligence (OSINT) feeds, social media or otherwise, which were the root cause of incorrect information propagation and amplification about COVID 19 vaccinations. Furthermore, we were attempting to understand the human networks, and their associations with specific groups responsible, noting the rise of extremist far right behaviours at this time, alongside the associated increases in hate speech. We observed, in my case at the same level of wonderment as the seismic data discovery, knowledge or network graphs visually representing communities across the globe actively sharing and amplifying messages which were simply untrue and conversations that were augmented with deepfake imagery. Machine learning algorithms developed within the software functionality were able to isolate and monitor the anomalous narratives created by, and propagated across, nefarious communities and in some case organisations.

The images, which we suspected were fake, or at least out of context, were deconstructed through machine learning, deconvolving them to a set of pixels, the pattern of which identified them as being generated by artificial intelligence. The pattern of the pixels is striking in AI generated images; they form a bright white spot in the centre and a scatter of RGB colour sequences radiating from this. They also leave a digital fingerprint of the technology which generated them, opening a new threat surface, whose vulnerabilities can be exploited by adversarial actors.

Just over one year later in 2022, a colleague from the UK Ministry of Defence (MoD) and I were working together, in another secure location, as Russia continued its aggressive attack on Ukraine. We had been tracking the increasing escalation since 2014, and the recent uptick in cyber and virtual activity directed towards Ukraine that was attributed to Russia.

Many in the defence and national security communities had been observing carefully since 2017 how Russian aggression had altered, shifting from activity in the physical, kinetic, and lethal domain, to that of the virtual and cognitive. Widely attributed to Russia, the introduction of the novel cyber pathogen, NotPetya, which leveraged the exploit EternalBlue, initially developed by the U.S. National Security Agency (NSA), caused local disruption to power stations and financial systems, and almost resulted in bankruptcy for the global shipping giant Maersk. NotPetya has been estimated to have caused over $10 Billion of global damage across eighty companies, as well as its intended target country, Ukraine. That is the central point here; it was an act of war against Ukraine, the consequences of which had a disproportionate effect globally.

"At what point do attributed, sub-threshold attacks spill over into an event which could plausibly escalate across the globe?"



An obvious and critically important question then arises: at what point do attributed, sub-threshold attacks spill over into an event which could plausibly escalate across the globe, triggering NATO's Article 5 mutual defence agreement?

The larger point is that from 2017 the nature of conflict was changing. Enabling technologies were making code ubiquitously available to state and non-state actors, and technology vulnerabilities, in this case software, were being exploited. The digital degradation of Ukraine was gaining pace, creating its own challenges at the beginning of the conflict in 2022. The weaponisation of data had begun.

My colleague and I had been tracking how the conflict in Ukraine had brought about greater awareness of the organisation Bellingcat. Bellingcat had been reporting the atrocities of the Russo-Ukrainian conflict using its powers of crowd-sourced investigative journalism. The organisation uses vast quantities of OSINT and machine intelligent processes, to correlate and fuse data sources to uncover the facts around war crimes and crimes against humanity, providing legal evidence that supports the deliberations of the UN Security Council and other international bodies.

Our work focussed on the methodologies Bellingcat had deployed to geolocate a photograph of the Russian Missile Programming Team purported to be behind programming many of the cruise missiles that hit Ukraine at the beginning of the conflict. Bellingcat had used multiple methods to identify the location of the photograph, and the details of those pictured in it, using various OSINT sources; satellite imagery, pattern of life, pattern of movement, facial recognition, image detection, all defined using machine intelligent processes. This was evidence which would later prove critical in defining war crimes.

How could we create, using similar techniques to Bellingcat, a capability which would support intelligence analysis for defence? We knew we had to develop a concept system which was able to harvest multiple, highly sophisticated, data sources, then apply data ontologies which would allow these sources to be integrated and, through machine intelligent processes, reveal multi-attributes from the physical and virtual domains, thus providing us with insight which would have otherwise remained invisible to the intelligence analyst. Our methodologies had to be auditable, leaving a digital trail of the code we had used to develop algorithms that interrogate the complexities and volumes of data we had to process, and analyse. Why? Because this was a multi-domain targeting tool, which would support human decision-making when recognising and engaging with assets of interest.

We posed our hypothesis in an approach paper entitled The British Army's Approach to AI which was published by the Ministry of Defence in December 2023. The critical enablers, outlined in the approach paper, have been brought to life in what is now a live project, building upon the concept system described previously.

The system builds on the principles I first deployed in 2002 and takes into consideration how the nature of conflict has altered and evolved. It fuses, blends, and cross-correlates relevant data sources, revealing hidden attributes through refined algorithms which interpret, interrogate, and query data to define successful targeting outcomes. These principles are underpinned by our approach to the application of data analytics, machine learning, and AIs against the MoD's ambitious, safe and responsible applications of AI doctrine, and a new directive known as Dependable Artificial Intelligence in Defence (Joint Service Publication 936 [draft]).

I mentioned earlier the need for the correct use of technical vocabulary from the get-go. As a defence and national security collective, we must increase our digital literacy, or employ expert practitioners who understand the deep science that lies behind the words used to describe and prescribe the desired military outcomes intended for the technology development. Too often I hear the wrong words used; the explicit use of the right language is critical in ensuring we create solutions which will truly compete with those of our adversaries, rather than falling short because incorrect terminology was used at the requirement and scope of work setting phase. Carl von Clausewitz, the nineteenth century philosopher-practitioner of war, once observed that the analysis of war can prompt an 'ostentatious exhibition of ideas.' A 'serious menace', he suggested, is the 'retinue of jargon, technicalities, and metaphors' that 'swarm everywhere – a lawless rabble of camp followers.'

Invisible Combat

Elizabeth Vallance-Bull

Since leaving oil and gas exploration and moving into defence and national security (NS), seven years ago, with the explicit direction to "bring your learning from the oil and gas sector to Defence and NS", I have observed the changing nature of contemporary war. I made this point earlier and would like to take the time to explain some of my observations.

The emergence of Cloud compute and the exploitation of the internet (dark web, and chat chambers, both of which have enabled the ubiquitous availability of OSINT) are the key differences between developing defence and NS capabilities now, and my work in 2002. Cloud has created ubiquitous availability of technology, democratised it, and created interconnectivity (and thereby the weaponisation of interdependence), which has further diluted the meaning of the state within a competitive international system.

Warfare is increasingly waged through surrogates. These are technological tools, and human capital, which absorb the patron's political, operational, or financial burden of conflict. The military are adapting emergent technologies, developing surrogates, and weaponizing them; data and artificial intelligence are examples of this.

Surrogates have the ability to disrupt the battlespace kinetically, the information space subversively, and the willpower of the adversary psychologically, all without a major combat operation. Surrogates scale rapidly, once code is released. They proliferate through ubiquitous and democratised access, horizontally (across states) and vertically (through state and non-state actors). Furthermore, emergent technologies can reinforce one another making them even more powerful.

Invisible Combat

Elizabeth Vallance-Bull

The principles of data-driven decision making, and the use of AIs for targeting, are not new. The examples from my early career and the clear, golden, data-driven decision-making thread interwoven through them demonstrates this. It also identifies that defence and NS is at best some 30 years behind adjacent industries, like oil and gas.

Autonomous weapon systems have tended to dominate the discussion about AI in military applications, no doubt highlighted by events in Ukraine. Less attention has been paid to the use of AI in systems which support human decisions in armed conflicts.

Tools such as the concept system being developed by the Army Futures AI Team, are known as AI Decision Support Systems (AI-DSS), which are a novel concept for defence. These are computerised tools which use AI software to display, synthesise, and/or analyse data, and in some cases make recommendations, even predictions, to aid human decision-making in war. These are not unlike the systems and software I was using in 2002 for seismic data mining.

The use of AI-DSS for targeting purposes, in all domains; physical, virtual, and cognitive, demonstrates how the traditional domains of air, land, sea, space, and now cyberspace, have evolved and converged. The consequence is that an ever-growing proportion of the contemporary battlefield bears little resemblance to the armoury of previous global conflicts, for which international legal frameworks were originally constructed. This presents us with moral, ethical and political challenges, as conflict moves from a purely human endeavour to a world of digital proliferation and data-driven decisions.

Get in touch!

Libby Vallance-Bull, libby.vallance-bull@capgemini.com