Viewpoint

Artificial Intelligence – Is the RAF ready?…Are you?

By Squadron Leader Andy Webb

Biography: Squadron Leader Andy Webb joined the RAF as a Fighter Controller, specialising in the human and technical systems which assure our surveillance of air and space. In 2015 he retrained as Aircrew, flying Reaper on Operation Shader. After a tour at the Military Aviation Authority, responsible for Aircrew Regulation, he has now joined No. 54 Squadron (the ISTAR[1] Academy) to lead Protector training. Andy is currently studying for a PhD, assessing how Artificial Intelligence (AI) affects organisational design.

Abstract: Continuous advancements in AI have led to the emergence of Large Language Models (LLMs) that utilise generative AI technology, and have the capability to effectively synthesise existing material, resulting in significant public interest. Increased interest has naturally led to substantial academic investigation, discussion, and investment. It is imperative to thoroughly analyse the ramifications of the current increase in AI for various organisations, such as the Royal Air Force (RAF). This viewpoint argues against the notion of AI being perceived as a comprehensive, all-powerful, or all-knowing solution, in contrast to the exaggerated tales that dominate conversations about its potential. This viewpoint will show that AI application requires an organisational plan for implementation to increase adaptation. The paper will cover AI preconditions and human-machine teaming to determine what modifications are needed, why RAF needs to change, and how the RAF as a human-focussed organisation may transition to an AI-enabled future.

Introduction

Artificial Intelligence (AI) is in vogue. The development of Large Language Models (LLM) derived from neural networks - which enable impressive articulation of pre-existing information - have captured the public imagination. Such popularity inevitably drives considerable study, comment, and investment. But what does that mean for organisations? What does it mean for the RAF? This short viewpoint argues that AI is not the apocryphal, omnipotent, omniscient answer-to-all-things that more ardent commentary may have us believe. Nor is it imminently a threat to our way of life. Instead, this paper will examine emergent research to demonstrate the need for an organisational strategy to unlock opportunities offered by AI applications. To assess what changes might be required, the paper will discuss some preconditions for effective AI and explore the need for effective human-machine teaming. This viewpoint suggests that why the RAF needs to adapt is evident from existing research, but that further research is urgently needed to define *how* we modify our existing human-focussed organisation to an increasingly AI-enabled future.

Previous editions of this journal have included multiple articles about AI. This is unsurprising – analysis of the distribution of academic papers with 'artificial intelligence' in the title show an exponential increase in recent years.[2] But our lexicography for AI is still developing. Referring to AI without adding context is unhelpful. It is akin to saying that this journal has featured articles about warfare, or papers about aircraft. This might be helpful in narrowing down the subject area a little but does not go far enough to be useful. Across various media, we see a wide spectrum of what is meant by AI. In *Air and Space Power Review (ASPR)* this breadth has ranged from: Wing Commander Ali Morton's philosophical strategic analysis of a potential revolution in military affairs;[3] to Squadron Leader Carolyn Swinney's technical primer explaining the history, future, and challenges of integrating the technology into our organisation;[4] and Professor Peter Lee's outline for an ethics framework.[5] This paper builds on the content of these papers and other academic literature to explore whether our people, as individuals and as teams, are ready to exploit the potential of AI. To do so, the paper will revisit how we categorise AI, discuss some strengths and weaknesses, and explore potential implications of their introduction into the human-led teams that make up our organisation.

AI Evolution

It is important to recognise that the recent emergence of LLM such as ChatGPT – whilst compelling – is not the first big breakthrough in AI. LLMs use neural networks borne from advances in machine learning mechanisms to allow reasoning based on language. This renders AI capable of passing Alan Turing's eponymous 'Turing Test', first suggested in 1950 – a test of a machine's ability to be mistaken for a human based on textual interaction.[6] But plenty of evolution has occurred through innovation over previous decades. An analysis by Lu *et al* into the distribution of academic papers mentioning AI in the title describes the evolution of AI in three key phases.[7]

The First Phase. They identify a first phase from 1956-1980. In this analysis, the 1956 Dartmouth Conference is the starting point for AI; closely followed by development of the first neural network 'Perceptron' in 1957.

The Second Phase. A second phase, the 'Industrialisation' phase, emerged in 1980 until 2000. In this phase, Japan invested $850m in AI (which, with inflation, would be $2.7bn today); multi-layer neural networks appeared; companies began widespread adoption of 'expert-systems'; and Deep Blue played chess well enough to beat the reigning human champion – Gary Kasparov.

An Explosion Phase. The 'explosion' phase since 2000 has seen an AI beat the 'Turing Test 2014', Watson and AlphaGo defeating the best human players, and Geoffrey Hinton proposed an AI deep learning training method for AI which did not require direct human supervision. Most recently, these deep learning capabilities have enabled the LLM behaviours that capture our imagination today.

Categorisations of AI - the importance of Data

The current state of the art in AI surrounds the increasingly compelling ability of Artificial Narrow Intelligence (ANI) to complete a given task (or set of tasks) well. Whilst some types of ANI have been around for many years – such as the AI for gaming (eg Deep Blue) - recent breakthroughs in machine learning and 'deep learning' neural networks have enabled the LLMs that are becoming increasingly well known in the public domain. These LLMs can consistently perform better than the average human in verbal and numeric reasoning, in addition to certain professional assessments. Whilst they are vulnerable to limitations such as hallucination (where incorrect responses are provided without any indication that the AI reasoning has produced inaccuracies), risk can be mitigated by the way in which a human interacts with the model – with guides on optimisation readily available and people selling their services in getting the best out of these systems. Another issue with LLMs is that they are poor at distinguishing false information from correct information in their databases because they are not grounded in real-world knowledge. This is a significant vulnerability for networked LLMs, as they could potentially be misled by misinformation. On the other hand applications of ANI which are not focussed on reasoning based on text have existed for longer and continue to proliferate as they develop. Such AI models have the potential to significantly change existing human-based systems – for example, creating a risk-index to enable a virtual command and control facility for the Arctic, or complex traffic management to allow aircraft to route point-to-point without sequencing in airways. Irrespective of type, ANI already exceeds human performance in many applications.

As the timeline in figure 1 shows, breakthroughs in AI performance have been enabled by machine learning algorithms since the 1980s, and deep learning methods since the 2010s. AI designed through a process of 'deep learning' include Convolutional Neural Networks (CNNs) which are critically dependent on the availability of data to allow their training cycle to be

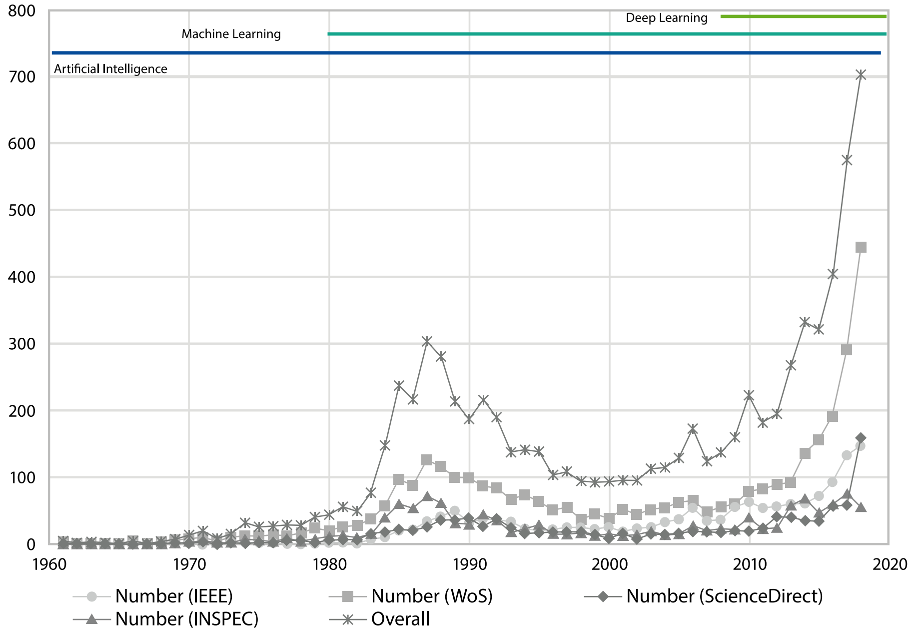

Figure 1 – Journal articles with 'Artificial Intelligence' in title (Lu *et al* (2019), overlayed with a broad characterisation of AI maturity (derived from Swinney (2022)).

effective[8] – although recent research is demonstrating that once a foundational model (such as GPT-4) has been produced, clever application of example data when training an AI might negate some of the need for 'tuning' in a particular use-case (few-shot training).[9] This reliance on data, and an understanding of how an AI is initially training and later tuned, is an important criticality for the RAF and for us. Key questions to scope future training for our people are:

- How many of us are confident in our contribution to effective 'Big Data'?
- What statistics training have we had to understand what data to capture, how to capture it in a useful way, and how to protect its reliability?
- Where do we draw the line between AI design, and AI use?

One analogy might be the knowledge gap between Aircrew flying an aircraft, the Engineers maintaining them, and the air and space operations professionals advising them. In this analogy data would be akin not only to the charts required for a specified mission, but also the combined knowledge that pilots and engineers were trained in and the fuel to power the aircraft. For AI, data is the essential component that supports all facets of the capability. This means data is both a key vulnerability and opportunity - the organisation needs to be configured to both protect and exploit data. There is consensus[10] that Defence is starting from a disadvantage – classified information, the spread of data across different systems, and a lack of consistent data (in format and in content) present 'wicked problems' that are foundational to gaining an advantage in AI development and integration. Finding solutions cannot simply be the purview of cyber professionals, as the experts in relevant data are those in each of the RAF's other professions - who gather, use, and employ data and information in their operational role.

But what about more sensational claims about AI? Given rapid progression in ANI capabilities and its increasing proliferation, it should be no surprise that there is much discussion in the field about the point at which a capability might be demonstrated that represents an Artificial General Intelligence (AGI). This is not simply the development of programming that can achieve multiple tasks with a single algorithm (this was achieved by the General Problem Solver in 1959).[11] Instead, AGI might be described as one which can achieve a fully representative cross-section of intellectual tasks that humans or animals can perform.[12] It is notable that the definition for what might constitute AGI has grown more challenging over time, as various instances of ANI have demonstrated increasing capability and flexibility. Estimates for when an AGI may be available vary wildly with some guessing 2029[13] and others the year 2030[14] – it is still a wholly theoretical capability which is ambiguously defined. By extrapolation of previous trends, we can expect that as ANI applications develop and become more flexible the behaviour that an AGI must demonstrate to be accepted as such will become more comprehensive. The hypothesised advent of an Artificial Super Intelligence (ASI), which would represent an AGI that surpasses the brightest human minds, is beyond the scope of this paper – as there is scant evidence it will be available during our lifetime. In the absence of a credible timeline for AGI and ASI, any of the effects posited as a 'Revolution in Military Affairs' by Wg Cdr Morton in *ASPR* last year should be examined in the context of the design-parameters for individual ANI and how they interact with existing human teams and organisational design.

The Human-Machine Team – a requirement for flexible levels of human input

Whilst an AGI may or may not be developed within our career-horizons, there is no doubt that ANI will continue to proliferate and increasingly affect all of us. The Ministry of Defence (MOD) has established centres of excellence to drive the strategic development of AI such as the Defence AI Centre. In turn, the RAF has established the Rapid Capability Office (RCO) with an Air Information Experimentation Laboratory (AIX)[15] which innovates with AI technologies. Some units of the RAF routinely use or develop AI in practical applications (such as the GUARDIAN Command and Control system, and 90 Signals Unit respectively). But there is less fidelity on how our broader systems, processes and people in the organisation may need to change – and limited evidence of an organisational strategy to enable it.

The RAF proudly declares itself to be a people-focussed organisation, with valuing people a key part of its overall organisational strategy.[16] I remember being told when I joined the organisation that our training, culture, and meritocracy ensured we were an Air Force that was, 'person for person, second to none'. This emphasis on people is a recognition of the centrality of the human in warfare, and however sophisticated our equipment and communication networks become, most decisions and judgements that result in success or failure have hitherto been made by a person. In the absence of a significant forcing-function (such as a peer conflict which threatens our way of life), humans are unlikely to be written out of the equation. Legal, moral, and ethical challenges[17] drive a policy of human-on-the-loop system design with 'meaningful human control',[18] to allow a sub-set of decisions to be reserved for a human. Unsurprisingly then, as existing UK doctrine explains, operational advantage is expected to lie with the most effective human-machine teams.[19]

Effective human-machine teams will be those that can most effectively leverage the respective advantages of human capabilities and automation. For example, the traditional role of Flight Engineer in aircraft flight decks was removed when it became clear that automation could create efficiencies that made a smaller crew complement credible - but full automation is usually deemed undesirable. At an organisational level, in very broad terms, automation allows for more rapid processing and (if extended to decision-making) action by fewer people. Higher degrees of human control often slow automated processes, but if well-managed could also make it more reliable and adaptable. In contested circumstances with an adversary seeking to disrupt our activity (by making decisions more quickly than we can) the degree to which people are involved in each process is likely to be, by necessity, highly situational. Figure 2 (below) illustrates how our preferred allocation of tasks between human and machine

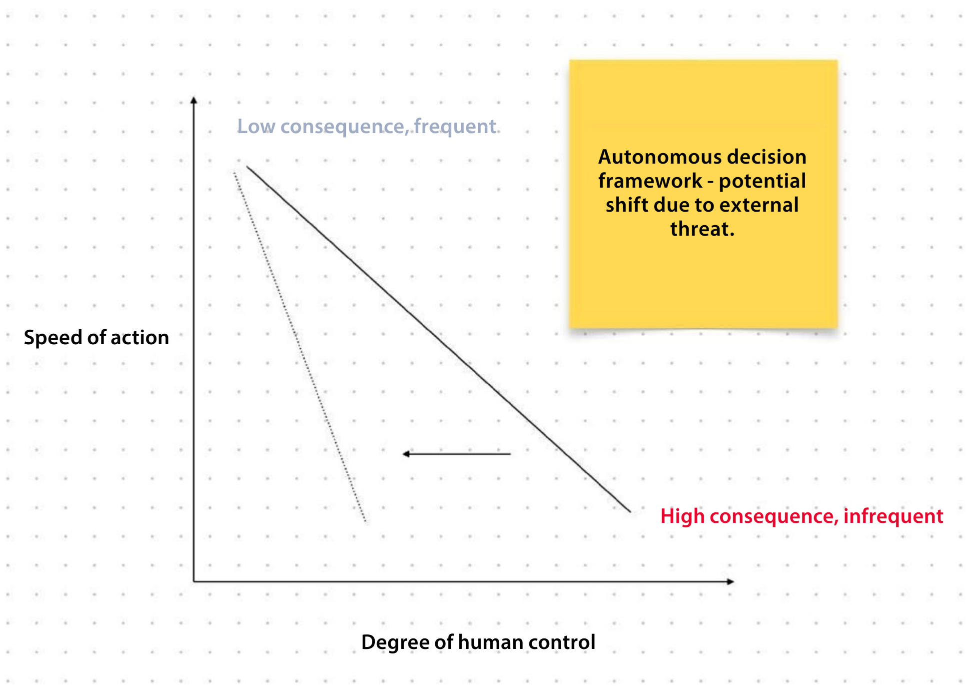

Figure 2 – Conceptualisation of AI autonomy change required as adversary decision cycle accelerates.

may need to shift, if constrained by the speed of an adversary's decision cycle (OODA loop).[20] The process and procedure to do this will need to be understood in advance to ensure appropriate decisions are made when circumstances become contested. Where AI systems support different parts of the organisation, this will also need to be coordinated to ensure appropriate alignment between the pace-of-decisions in complementary systems. There will be no point in a lightning-fast supply allocation tool which is constrained by the availability of human movers or a slow system of prioritisation due to misalignment with decision making at an operational headquarters. In turn, this requirement for coherence is likely to require increased mutual understanding of related parts of the organisation – requiring an interdisciplinary approach between the RAF's newly redefined Professions. Furthermore, AI models need to be designed to cater for these shifts and human teams need to be sufficiently informed to accommodate them. This will require familiarity with AI concepts, which do not currently feature in generic professional military education.[21] Therefore, development of a suitable taxonomy to familiarise RAF personnel with AI must be a priority.

The Human-Machine Team – the transparency paradox

As Swinney explained in her 2022 article, two key technical challenges for AI are bias and variance. These are also human frailties. One way in which human bias overlaps with AI capability is in the extent to which a person is inclined to accept the recommendation or decisions of an AI. Our decisions as humans are conditioned by training, circumstance, and the broader organisational system to which we are exposed. So far, so similar to AI bias. But a key mitigation for human bias is the ability to perceive a pre-disposition due to our understanding of perspective and adjust accordingly. The extent to which an operator can do this in relation to an AI will depend on a knowledge of its training data and an assessment of likely bias or variance in the data. Counterintuitively, there is recent evidence that people are less likely accept a decision from an AI when they know more about how it works (due to overconfidence) and that this can lead to poorer outcomes.[22] It will also be affected by the extent to which an AI is subject to 'social proof' – by including respected peers in development and testing. The siloed nature of hierarchical organisations like the military, and the 'need to know' ethos that underpins classified information management, render teams necessarily more likely to accept 'on faith' assertions that pop out of another bit of the organisation.[23] Without systemic intervention, knowledge of AI design is likely to be held only by those who design the AI – and simply upskilling people may not be the answer – especially if further research provides additional evidence that people tend to be overconfident in overruling AI output. Instead, a systemic approach may need to be defined as part of through-career generic training. Such a systemic approach would need to balance sufficient transparency to allow the flexibility required of human-machine teaming to stay inside an adversary's decision cycle, without allowing for exclusion of AI input because 'I know better' or total abdication of human responsibility. Fundamental to it all would be an ability to understand the importance of, and judge data quality.

Developing a Strategy – estimating workforce changes

If we accept the premise that AI poses a need for us to adapt our human teams (*why*), then the next question is *how*. A logical starting point for a strategy is the front door – who we recruit.

Recent research into the impact of AI on the labour market suggests AI will have a transformative effect on the higher-skilled workforce.[24] Analysis of job task descriptions and AI patents suggests that, in contrast to robots and software, AI is directed at high-skilled tasks. If the historical pattern of new technologies substituting for a proportion of human jobs continues, then unlike robots (which disproportionately displaced low-skilled human labour) and software (which most impacted the middle-skilled), AI will displace a greater proportion of the higher skilled.

If we place this in an RAF context, it poses some interesting possibilities. It is likely that this trend could, over time, reduce the proportion of senior officers or experienced specialists within the workforce. Some evidence of this has been seen in the recent Royal Navy Transformation programme which reduced by up to 30% senior ranks and positions. More research could help the Service to plan for a changing workforce, by assessing further demographic variables, including occupational salary (Recruitment and Retention payments), educational level, gender and age. It may also be possible (using an adapted research methodology) to use the Training Needs Analysis (TNA) conducted for each RAF job type under the Defence Systems Approach to Training[25] to further refine such an estimate. Where a

Figure 3 – Approximation of the relationship between salary / skills and exposure of jobs to substitution by robots, software, and AI (derived from Webb (2019)).

TNA is available for both a role in a non-AI system and one which is later AI-enabled, this could also enable the RAF AI strategy to explore and define any delta between optimal (peacetime) and degraded (wartime resilience) requirements.

Developing a Strategy – options for training

If adaption of our workforce is one element of our AI strategy, another must be evolution of our training. In the RAF training is often described as professional (specific to a career field) or generic (applicable to all in Service).

Training is further categorised by Phase – Phase 1 (Basic Recruit or Initial Officer Training), Phase 2 (the first professional training resulting in a certificate of competence in role), and Phase 3 (continued professional development). This framework could help to situate training opportunities for AI-literacy in a familiar context.

One training opportunity is to ensure we have people confident in codifying AI systems in a contextual taxonomy – which is crucial in an ethical context and therefore fundamental to the moral component of warfare. Emergent research into the ethics of development and deployment of AI systems in high-risk domains has emphasised the importance of an established methodology for interpreting AI ethics principles in a way that is consistent with the context in which they are used.[26] It is reasonable to expect that the need to rapidly procure new systems during a conflict will require familiarity with this concept. The starting point for training these skills is likely to be familiarity with ethical threats and mitigations, together with fluency in the way the AI system works. In turn, this will require development of appropriate categorisations which will drive a broader taxonomy for their use and level of human control. The practical application of the training would be attribution of a category to a given AI use case and identification of any operating constraints. This management of ethical risk (impact and likelihood) is likely to be necessarily limited to specialists and therefore be a component of professional training – with contributions likely from the Cyber, Chaplaincy and Operations Professions. Given that the degree of exposure to this task will vary by assignment, this might be best suited to Phase 3 training.

More generic training may help aviators to understand the way that AI operates at the level of the system in which AI is an agent. 'Systems thinking' is a way of assessing behaviours of a system (and how to influence changes) which may not be evident if we simply look at the component parts. This is likely to have significant advantages for understanding the relationships between human teams that are interacting with AI. As a generic skill, it also offers leaders the opportunity to improve the performance of their teams whether AI is involved or not. Fundamentally, it offers a framework for understanding complex processes.

When we try to understand something, we traditionally engage in reductionism - where we seek to identify all the components of a problem, take them apart to understand them, draw conclusions and then improve them. In contrast, systems thinking is more interested in relationships and outcomes when the system itself is observed as a functioning entity.[27] The theory of systems thinking has emerged from biology, where some characteristics of biological systems can only be observed at a systemic level. Anatomic knowledge of an animal does not allow understanding of its behaviour, nor for the interaction between organs for example. This is particularly relevant to AI, as neural networks are not fully transparent and therefore, we are likely to learn more through its behaviour than simply its training. Furthermore, understanding the AI would not be enough – as we would also need to understand the behaviour of people who work with the AI, and the way in which that affects the overall organisation or domain. In a systems-thinking approach it is not the AI thought process itself that would be examined, but the overall behaviour of the wider system in which it acts as an agent.

Systems-thinking is therefore likely to be a desirable competence for all future aviators - but is currently unfamiliar to most. It might be that this renders it particularly suitable to graduated training which is re-introduced and reinforced in new contexts as people progress throughout their careers. A generic training pathway would provide a perfect opportunity to define the requirement for inclusion in Phase 1, 2 and 3 training. In this way the techniques and models could be introduced in ways that were directly relevant to individuals at that point in their career. Together with data literacy, it could be a worthy addition to generic professional military education. Alternatively, an approach similar to human factors training might be taken – where a Regulation or policy drives a requirement of levels of competence in dealing with errors borne of human factors in human teams or when working with automation.[28]

Existing Subjects[29]	**Potential Subjects**
Air and Space Power	Air and Space Power
Leadership	Command, Leadership and Management[30]
Management	Data Literacy
General Service Knowledge and Skills (GSK)	Systems Thinking
Command	GSK (inc Ethics, International Humanitarian Law)
Force Protection	Force Protection

Figure 4 – Table comparing current / potential future subjects for generic professional military education.

Conclusion

This viewpoint has offered a précis of research which makes it evident that the RAF will need to change and evolve as AI proliferates. The first step to shared and effective understanding is to communicate the distinction between AI types and be clear-eyed about the capabilities of each instance of AI in the organisation. Given that AI is reliant on effective training, and in turn data, the requirement for improved data literacy is urgent. We don't know what level of human input is required in each future application of AI, but evidence shows us that the level of input is likely to be contextual – and potentially constrained by an adversary or competitor. A means of considering the ethical implications of a given system given its use-case, data-training and design is likely to be an essential contributor to determining the level of human input – and therefore require a professional training pathway. Even when a system is accepted into service and employed, we need to better understand the human factors – just as we do for human teams and Aircrew working with automation. As AI proliferates it will affect the labour market over time – and this might drive changes in the RAF workforce strategy. Development of generic skills including systems-thinking might drive the inter-disciplinary dialogue and a better understanding of AI models influence within the organisation.

Notes

[1] ISTAR – Intelligence, Surveillance, Target Acquisition and Reconnaissance.
[2] Lu (2019)
[3] Morton (2023)
[4] Swinney (2022)
[5] Lee (2019)
[6] Turing (1950)
[7] Lu (2019)
[8] Alzubadi *et al* (2021).
[9] Vinyals *et al* (2016); Sucholutsky and Schonlau (2021).
[10] A common theme throughout keynote speeches at AI Fest 5 – DSTL (2023).
[11] Newell and Simon (1959) as cited by Swinney (2022).
[12] Shevlin *et al* (2019).
[13] Kurzwell - Director of Engineering at Google, as cited by Swinney (2022).
[14] Brooks - co-founder at Robust.ai (2019).
[15] Wigston (2020)
[16] RAF (2022).
[17] Vohs (2021).
[18] UN Office for Disarmament Affairs – Convention on Certain Conventional Weapons (2023).
[19] Joint Concept Note 1/18.
[20] As developed from the ideas of Boyd, see Joint Concept Note 1/20 – Multi-Domain Integration.
[21] AP 7000 v3.1 (2021).
[22] DeStefano *et al* (2022).
[23] Dawes, Cresswell and Pardo (2009).
[24] Webb (2019)
[25] JSP 822
[26] Taddeo *et al* (2023)
[27] Meadows (2008)
[28] RA 1440 and MAA Human Factors Training Requirements.
[29] AP 7000 v3.1 (2021)

[30] In delivery, these separate elements from the generic performance statement are *de facto* delivered together at a level set by rank.