Data quality and analysts' role in AI enhanced C2

Erik Bjurström	Martin Schüler	Anette Strömberg	Git Roxström
Mälardalen Unversity	*Unversity West*	*Mälardalen University*	*Swedish Defence Research Agency*
erik.bjurstrom@mdu.se	martin.schuler@hv.se	anette.stromberg@mdu.se	git.roxstrom@foi.se

Abstract

Artificial Intelligence (AI) in Command and Control (C2) raises questions about the interaction between operators on different levels on the one hand and AI-supported information systems on the other. For some purposes, the aggregation and analysis of large "big data"-sets creates potential for prediction and calculations of probabilities, while for other purposes human heuristics may be as promising depending on the situation. What has been more rarely discussed, is the very quality of data underpinning such calculations, and thus also operators' awareness of the validity of predictions in relation to any specific situation. This is an urgent debate, considering the fact that full transparency may be impossible and underpinning data may be based on exercises, simulations, real-time data, or a mix thereof. Based on a fundamental classification of different kinds of uncertainties, this article discusses how data of different origins and quality can be managed and communicated to allow for operators to assess on what basis predictions are made. The article further suggests that looking at other fields of research may be useful for exploring unconventional ways of highlighting the existence and quality of different kinds of data of different origins in order to assess its predictive power. Finally, the article discusses how AI may change the role of analysts with regard to such issues.

1 Introduction

Decision-making is at the heart of the military profession. When asked about their central task, many officers reply exactly that: – To take decisions. While for professional decision-makers, 'less than AI-capabilities' have been embedded in information systems for a long time, influencing access, distribution and presentation of information, public awareness of such technologies suddenly changed. Artificial Intelligence (AI), that for decades has been talked about, yet seeming somewhat distant, visionary and hypothetical, all of a sudden became present at everybody's fingertips through the advent of Chat GPT in the late autumn of 2022. This new reality sent chock-waves through the public sphere, for politicians, policymakers and in academia, raising questions about academic examination, integrity of citizens (through GDPR), transparency of algorithms (explainable AI), challenges to administrative work (summarizing knowledge; writing reports), and for the general public wondering what the personal and societal consequences might be. Prominent intellectuals, researchers and celebrities called for a ban of further development of a technology that presently is capable of going from being an absolute beginner to an unbeatable chess champion in only on day, through the mere speed of iterations.

While AI – through its broad definition as a technology doing what hereto only humans could do [1] – has by tradition been around in military equipment far earlier than in civil applications generally, this latest technological breakthrough more reminds of the chock that was caused by the advent of Play Station II, which was the first toy that had to pass through military technology quarantine before being introduced to the market in 1999, because of its powerful graphics processor [2]. This event was chocking but should have been expected as globalization generated a demand and funding for R&D for the entertainment market that for the first time could compete even with military spending on development of advanced technology.

Reactions on Chat GPT have so far been moderate in the military field but has the potential of more far-reaching consequences than Play Station II ever had, as it touches at the heart of the very notions of excellence, not only in military technology, but of the military profession itself. Indeed, the slow reaction to this potential provocation – at least in Europe – may reflect the "transatlantic gap" in the development of software-based defense [3]. With the speed by which new versions of Chat GPT has developed during the last year, there is already reason to ask what the effects may be on one of the hereto rewarding efforts in terms of citations, i.e. literature reviews, both on the production side, compiling established research, and consequently also on the consumption side, through their decreasing value, as Chat GPT increasingly able to provide quick overviews over new topics. In a similar vein, military staffs' capacity to compile relevant information from a broad array of relevant sources is increasingly being challenged by both the accuracy and speed of automated generation of overviews, including suggestions for alternative courses of action. This has even greater potential for impact in the military profession, given the emphasis on speed, i.e. to get inside an adversary's decision-loop.

2 AI and heuristics

While insistence on the speed of decision-making may be questioned [4], it has typically been the trump card in the argument for human heuristics. Simple ranking of alternatives may be as efficient, and even superior to, attempts at optimization in an uncertain world where speed and simplicity does matter [5]. However, with increasing calculative power comes also further speed and the possibility to obtain both local, situated accuracy and – at the same time – coordination and possibly also direction, i.e. command and control (C2) in terms of its *function* [6]. It is predictable that AI systems will increasingly become an integral part of the C2 system in a software-based defence. Consequently, C2 as a matter of *design* [7, 8] will also to a greater extent be a matter of viewing human and 'non-human partners' [9] as integral components of the C2 system of a software-based defence.

This development raises all kinds of questions about operators' and military staff members' trust in their non-human partners. In consequence, there has been an explosion of interest in issues of explainable AI not least from DARPA [9] and the political leadership with the vision of protecting citizen's integrity through e.g. the European Union's General Data Protection Regulation (GDPR) and a demand for transparency of AI algorithms. From an operator's or staff member's perspective, it becomes more a question of who's judgment that should be trusted: the AI-driven calculation of probabilities, or the operator's or decision-makers of higher ranks' heuristic in any given situation. On a practical, perhaps somewhat superficial, level this choice may seem just a matter of making the better choice. However, considering the highly emotional aspects of the military ethos, human decisions may be preferred, whether more accurate or not, as many decisions may have consequences of life or death, hence being fundamentally social and ethical [10]. In a pragmatist sense, this affects the very notion of validity, which should include ethical considerations [11], which may seem to limit the range of choices available to decision-makers and the way the C2 system may be designed. On the other hand, a more positivist 'culture of objectivity' has been found to be associated with an over-confidence in quantitative data, leading to an insensitivity to uncertainty, ultimately resulting in an inaccurate estimation of risk [12]. Hence, the increasing influence of AI as a non-human partner in C2 calls for a deepened discussion about risk, uncertainties and how they are perceived and treaded – also as a possible threat to the core of the military profession's task as decision-maker, as well as the professional ethos, identity and ethics.

3 Classifications of uncertainties

Uncertainty and how it is addressed has not only an intellectual side to it (how we think about it), but also a practical side (how we actually treat it), and – just as importantly – an emotional side (how we feel when facing the issue) [13]. In the below 'uncertainty triangle', there are three extreme points, each representing 'primitive attitudes' to uncertainty. Immature actors tend to be drawn to such extremes, while maturation in the attitude includes the ability to consciously use combined approaches, i.e. playing over the entire field depending on circumstances [13, 14].

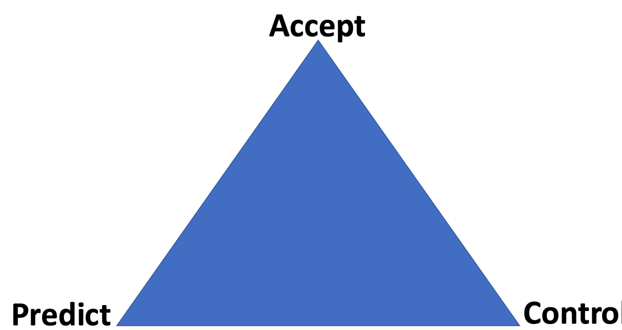

Figure 1: 'The uncertainty triangle' (Eriksson, E. A., p.169)

Apart from these extreme positions in the basic attitude towards uncertainty, and the maturation to a more balanced approach, there are a number of fundamental aspects of uncertainty that need to be addressed to problematize and nuance the discussion about what the shifting balance between humans and non-humans in an integrated C2 system may imply.

3.1 Quantitative vs. qualitative uncertainty

A basic distinction is that between quantitative and qualitative uncertainties. Most academic work takes the probabilistic view on uncertainties, hence implying quantitative uncertainty, which is appropriate when future events can be expected to obey the same fundamental logic. On the other hand, qualitative uncertainty refers to a situation where alternative future events obey structurally different logics [14]. Ideally, for quantitative uncertainty to being relevant, we need a closed probability space, and historical frequencies, to know the expected distribution of future events within that space. While Bayesian probability include subjective estimates of that distribution, it still follows the attitude to uncertainty follows the same basic logic. Qualitative uncertainty on the other hand not only means that we don't know how often an event will happen, i.e. its distribution in terms of probabilities, but also not *what* may happen, i.e. that there is no closed space within which probabilities can be distributed. Hence, the notions of quantitative vs. qualitative uncertainty will be fundamental for an operator's assessment of whom to trust in the face of uncertainty.

3.2 Intentional vs. stochastic uncertainty

A second distinction of fundamental importance for the design of any combination and balance between human and non-human actors in the C2 system is the basic question: - What are we up against? [14] Being faced with a logistical problem, requiring only a coordinated C2 approach, should be distinguished from a situation where fluctuations and events are not of stochastic character, i.e. randomly distributed, but rather an effect of intentionality of, still predominantly, human actors. While the frequent use of the notion of 'autonomous' non-human actors, there is still a distinct difference between human autonomy and that of a self-driving vehicle, precisely in the sense human autonomy builds on notions of intentionality of human action, such as a narrative mode of cognition [15, 16]. Hence, when facing uncertainty in a real-life situation – especially in conflict – the very character of uncertainty is of a fundamentally different kind, even when calculating probabilities or ranking alternatives for action through more or less intuitive heuristics, more similar to interpretative research than quantitative analysis in that it is *all* about emphatic interpretation of what is going on in an adversary's mind; what story that is guiding his intentions and action, which is fundamental for warfare's longstanding practices of deception.

3.3 Dynamic vs. static uncertainty

A third fundamental distinction is that between static uncertainty [14], where a decision-maker is more or less forced to make a high-stake decision under uncertainty, without any greater room for later adjustment or adaptation on the one hand, and dynamic uncertainty where the situation allows for probing actions. While much research into decision-making experiments have followed the first line of thought of Kahneman & Twersky's 'prospect theory' [17], while research into 'dynamic decision-making' [18] has gained less attention in general, while being on the other hand perhaps even exaggerated through notions of maneuver warfare in an era of AI based on real-time data. Technically speaking, static uncertainty is about the relation between warning time and response time: where response time is longer than warning time, uncertainties are static, while where response time is shorter than warning time, uncertainties are dynamic. Consequently, shortening of reaction time may convert static uncertainty into dynamic uncertainty. However, in the face of AI systems, this core of military common sense about 'getting inside the adversaries decision-loop' is severely challenged and may imply near total irrelevance of maneuver warfare for any actor that hasn't integrated non-human AI partners in their design of the C2 system. Ironically, the remaining remedies may only be massive force, attrition warfare at high costs, and exploitation of intentionality through deception.

4 Classification of data feeding the AI

Yet another, and often overlooked, aspect of AI is the question of by which data it has been trained – beyond notions of supervised or unsupervised training. Despite the massive resources spent on both on military research on the one hand, and training and exercises on the other hand, only a small fraction of the data generated is typically transferred between these spheres and used for development of knowledge and practice. Instead, large-scale research programs with military funding run in parallel with, and still in splendid isolation from, military practice. In fact, coordinating different streams of research seems big enough of a challenge [9]. Listing their top 10 recommendation [20] emphasized the importance of small-scale experiments for AI development, with real-world operators actually having knowledge about their task and equipment, in order to identify any gains of the new technology before scaling up to massive research programs based on speculative hypotheses.

It may seem strange that this situation prevails despite the ongoing AI race [3], primarily between the U.S. and China, while Europe is evidently lagging behind on the other side of the 'transatlantic gap'. The reason for this may, however, lie in the historic traditions of operations analysis and the divide between 'hard' and 'soft' operations analysis, beginning in the 1950s and becoming obvious in the 1970s, with a more quantitative (predominantly American) stream, and a more qualitative (predominantly British) stream of operations analysis [19]. However, and ironically, while this divide has gotten a fair amount of attention within the operations analysis community, the attention to practice has not. Instead, the call for more practice-based exploration of possibilities for AI enhancement [20] is practically not met by any enthusiasm by the analyst community. This has profound consequences for the development of AI through the kind of data that is used for its training.

From an operator's perspective it makes a world of a difference to know whether the non-human AI-partners' recommendations is based on:

- assumptions-based simulations,
- one-sided exercise data,
- double-sided exercise data,
- real-time combat data

… as well as the proportions thereof. The assessment of the quality of data as well as of its relevance for training the non-human AI partner will be of outmost importance, when weighing the value of its contributions against the core of the professional ethos, rather leaning towards the emotions aspect of uncertainty, rather accepting than predicting and controlling, and emphasizing the social and ethical – if not holy [10] – aspects of risking one's life as well as the life of others by accepting the responsibility of taking the decision. The severity of the trade-off between rationality and the military ethos [10] also makes it important to know the very quality of the data feeding and thereby forming the non-human partner, thus taking us back to the core of the most human of uncertainties; that of intentionality and the difference in quality between situations in exercises where fighting becomes as real as real-world combat on the one hand, and exercises where participants evidently look at events with more distant calculations about how this or that action may influence their assessment and thus further career opportunities – even at the cost of violating the rules of the exercise at high risk [21]. Hence, not only the form of the data feeding the AI matters, but also the quality of data itself, in turn highlighting the issue of human intentionality as the gold-standard for data quality. If military ethos and professional norms are going to be challenged by non-human partners – this had better be good.

5 Looking at other fields

What emerges is a picture of the development of AI-enhanced capabilities of the C2 system that is ripe with challenges and contradictions, both technically and professionally. Emotional, as well as practical and intellectual aspects of how to deal with uncertainties – through prediction, control or acceptance – interact with different kind of uncertainties; quantitative, stochastic and static (thus needing calculative support to operators) on the one hand, and qualitative, intentional and dynamic on the other. In addition, there is a broad array of different kinds of data that may feed the AI-system, without any guarantee of the validity of rigorous models for simulation, realism of exercises and the subsequent mirroring of realities underpinning AI algorithms that should be balanced with the military ethos and culture. In addition, there is no way for an operator to validate the authenticity of the underlying data than through the mere performance of the non-human partner.

Looking for inspiration for how to deal with this challenge, we need to consider the complexity of the task, as well as the need for broad societal data from different sources, while not being able to standardize it, neither in definition, nor in amount or quality. One of few professions with such broad challenges can however be found in the domain of public health: far beyond the laboratory or even the clinical practice, public health has to deal with the entire breadth of human behavior in order to estimate what measures that would be adequate and what recommendations that would also be followed, given the insight that the ones carrying the ultimate risk should also have a say. In this domain, Topological Data Analysis has been suggested as a means to provide a graphical overview of the data available for different fields of relevance, hence providing an understanding of the basis for analysis, rather than looking to explain the analysis (in this case AI algorithms) itself. This approach might also have potential for the defence sector.

6 On the relevance of future operations research

It is evident that the era of automated analysis through AI presents a challenge not only to operators and military decision-makers, but also – and not least – to operations analysts. Ormerod raised the question of the fitness and survival of the operations research (OR) profession, commenting that 'the clock is ticking' and concluding with four good reasons for taking the AI threat to OR seriously [22]:

- There is an urgent need to develop an understanding of the AI phenomenon
- There is an opportunity to get more deeply involved with the development of AI

- The C2 research community, as well as the OR community needs to position itself in relation to AI
- The OR community needs to get started as it will take years before we see necessary changes in perceptions and habits involved

In line with the top ten recommendations for AI development [20], we suggest that such initiatives are taken in close collaboration with actual operators and their challenge to balance the character of their new non-human partners in the C2 system with the very character of military work.

References

[1] Rich, E. Artificial intelligence. McGraw-Hill, New York, NY (1983)

[2] Jensen, K. T. 20 years later: How concerns about weaponized consoles almost sunk the PS2. PC Magazine. May 9. 2020. Retrieved June 1, 2023 from:https://www.pcmag.com/news/20-years-later-how-concerns-about-weaponized-consoles-almost-sunk-the-ps2

[3] Soare, R.S., Singh, P., & Nouwens, M. Software-defined defence: Algorithms at war. London, UK: The International Institute for Strategic Studies. February, 2023.

[4] Schüler, M. & Bjurström, E. Speed vs. thought. Paper presented at the 26th International Command and Control Research and Technology Symposium, Oct 18-22; 25-28 (virtual). 2021.

[5] Katsikopolous, K. V., Simsek, Ö., Buckmann, M., & Gigerenzer. G. Classifciation in the wild. Cambridge, MA: The MIT Press. 2020

[6] Brehmer, B. Understanding the functions of C2 is the key to progress. International C2 Journal, 1(1). 2007.

[7] Brehmer, B. Command and control as design. Paper presented at the 15th International Command and Control Research and Technology Symposium, June 22-24, Santa Monica, CA. 2010.

[8] Boland, R. & Collopy, F. (Eds.) Managing as designing. Stanford, CA: Stanford University Press. 2004.

[9] Gunning, D, Vorm, E., Yunyan Wang, J., Turek, M. DARPA's explainable AI (XAI) program: A retrospective. Applied AI Letters 2(4), 1-11 (2021).

[10] Brodrej, P. (forthcoming). Potential AI application in the Swedish Army and its relation to the holiness of life: A study of army officers' narratives on the military experience in the face of AI. Master Thesis in Leadership & Command and control. Swedish Defence University.

[11] Putnam, H. (2004). The collapse of the fact/value dichotomy and other essays. Harvard University Press.

[12] Feldman, S. P. (2004). The culture of objectivity: Quantification, uncertainty, and the evaluation of risk at NASA. *Human Relations.* 57(6):691-718.

[13] Dreborg, K.-H., Eriksson, E. A., Jeppsson, U., & Jungmar, M. (1994). Planera för det okända? Om hantering av osäkerhet. (Planning for the unknown? Coping with uncertainty.) FOA report R—94-00005-1.2—SE. Stockholm FOA.

[14] Eriksson, E. A. (2004). Scenario-Based Methodologies for Strategy Development and Management of Change. In: Olsson, MO., Sjöstedt, G. (eds) Systems Approaches and Their Application. Springer, Dordrecht. pp. 167-194. https://doi.org/10.1007/1-4020-2370-7_7

[15] Czarniawska, B. Narratives in social science research. London: Sage. 2004.

[16] Bruner, J. Actual minds and possible worlds. Cambridge, MA: Harvard University Press.

[17] Kahneman, D. & Twersky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*. 47, 263-291.

[18] Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 262–279). University of Chicago Press.

[19] Frelin, J., Eriksson, T., Räty, R., Kindvall, G., Nordstrand, E., Lindberg, A., Reimer, H. Förstudie forsknings- och teknikutvecklingsområde militär operationsanalys (FoT OA). [Prestudy research and technology development area military operations analysis, in Swe-dish] FOI--R--4820—SE. Swedish Defence Research Establishment (FOI). Stockholm. (2020).

[20] Hoffman, R.R., Jalaeian, M., Kein, G., Jentsch, F., Clancey, W.J., Mueller, S.T. Top ten recommendations for the development and assessment of AI systems. Conference paper pre-sented at the 27th International Command and Control Research and Technology Sympo-sium. 25-27 October. Quebec City, Canada. (2022).

[21] Schüler, M. (2022). Planning for safety when preparing for war: Competition brings out the wickedness of the military. Paper presented at the 27th International Command and Control Research and Technology Symposium. Quebec City, Canada. 25-27 October. 2022.

[22] Richard Ormerod (2021) The fitness and survival of the OR profession in the age of artificial intelligence, Journal of the Operational Research Society, 72:1, 4-22, DOI 10.1080/01605682.2019.1650619