

Artificial Intelligence in Wargaming

An evidence-based assessment of AI applications

Anna Knack and Rosamund Powell

June 2023

About CETaS.....	4
Acknowledgements.....	4
Executive Summary.....	5
Recommendations.....	8
1. Introduction	9
1.1 Research aims and methodology.....	11
1.2 Caveats and limitations.....	12
1.3 Structure of this report.....	13
2. Progress Towards Implementing AI-enabled Wargaming	14
2.1 A renewed interest in automated wargaming.....	14
2.2 Case studies in AI-enabled wargaming	17
2.3 Opportunities in non-defence applications of AI.....	21
3. Promising Applications of AI in Wargaming.....	26
3.1 AI in narrow and specialised applications.....	26
3.2 High-risk, high-reward AI applications in game execution and data analysis	28
4. Technical Challenges and Requirements	31
4.1 Development of a valid methodology.....	31
4.2. Over-trust.....	32
4.3. Ensuring trustworthiness.....	33
4.4. Establishing causality	34
4.5. AI with contextual awareness.....	34
4.6. Data and AI system sovereignty.....	35
5. Policy and Ethical Challenges	36
5.1. Accountability and liability	36
5.2. Reduction of decision-making pressure on players.....	36
5.3. Psychologically preparing players for human-machine teaming in wargaming.....	37
5.4. Efficient defence procurement and approvals processes	37

6. Conclusions and Recommendations	39
Recommendations.....	39
Appendix A. Long-list of AI-enabled Techniques.....	41
A.1 Game design	42
A.2 Game execution.....	45
A.3 Game analysis	49
A.4 Game logistics.....	52
Appendix B. Glossary	54
B.1 Wargaming lexicon.....	54
B.2 AI lexicon.....	56
About the Authors	58

About CETaS

The Centre for Emerging Technology and Security (CETaS) is a research centre based at The Alan Turing Institute, the UK's national institute for data science and artificial intelligence. The Centre's mission is to inform UK security policy through evidence-based, interdisciplinary research on emerging technology issues. Connect with CETaS at cetas.turing.ac.uk.

This research was funded by the UK Defence Science and Technology Laboratory (Dstl) through The Alan Turing Institute's AI Research Centre for Defence (ARC-D). All views expressed in this report are those of the authors, and do not necessarily represent the views of The Alan Turing Institute or any other organisation.

Acknowledgements

The authors would like to thank all those who contributed their time to participate in an interview or workshop for this project, without whom the research would not have been possible. The authors are also very grateful to Lt Col. Scotty Black and Aaron Frank, for their valuable feedback on an earlier draft of this report.

Executive Summary

This report presents findings from CETaS research undertaken on behalf of the Dstl-sponsored AI Research Centre for Defence (ARC-D), examining the potential for the application of artificial intelligence (AI) and automation in wargaming. The research focused specifically on segments of manual analytic wargames with partially simulated elements. The study aimed to identify ways that AI could 1) increase the efficiency of preparing and implementing a game, 2) support player decision-making and 3) improve the insights that can be gained through wargames.

The findings are based on a literature review, expert interviews, case study analysis, and a workshop convening experts across the defence and game AI community. The study explored case studies trialling the integration of AI in wargaming such as Group W's BrainSTORM, DARPA's Gamebreaker AI, and DARPA's Stabilizing Hostilities through Arbitration and Diplomatic Engagement (SHADE) project. The study also examined non-defence applications of AI in commercial games, healthcare, and diplomacy, which may yield inspiration for segments of wargaming.

Wargaming and simulation continue to be crucial tools for decision-makers in Defence. They can be used to train personnel for future conflict, and offer insights on critical decisions in warfighting, peace negotiations, arms control, and emergency response. Nevertheless, wargames also come with several challenges including the length of time required to prepare games, high set up costs, limited repeatability, and incomplete data capture within wargames.

Given advances in non-defence AI, and game AI in particular, there is growing interest in leveraging AI for wargaming and simulation. The envisaged benefits are specific to the context of use, but examples include reducing the number of personnel required, increasing the speed of development of game mechanics, improving player immersion, speeding up game execution, and identifying innovative strategies and actions.

The research identified two key features of the current landscape of AI-enabled wargaming that make it harder for decision-makers to determine whether AI can in fact achieve these benefits:

1. This is a **nascent debate**, which has been heavily influenced by **AI hype**. While many ideas are circulating on how AI *could* improve wargaming workflows, few real-world case studies offer concrete evidence of effectiveness.

2. AI-enabled wargaming is a subject that prompts **highly varied opinions between subject matter experts**. Much disagreement can be attributed to differences in expertise and experience, for example between wargamers and experts in modelling and simulation, or between technical experts and strategic decision-makers.

Within this context, this report aims to advance the debate by taking an evidence-based approach to assessing the feasibility of specific AI use cases for wargaming, outlining both their risks and potential benefits.

A long list of promising use cases identified is presented in the below table, with each use case described and analysed in more detail in Appendix A.

Game design	Game execution	Game analysis	Game logistics
<ul style="list-style-type: none"> AI red team agents Procedural Content Generation Scenario and inject generation Gameplay testing 	<ul style="list-style-type: none"> Simulating opponents' responses Adjudication AI to run through tactical decisions AI decision assistants and Course of Action analysis tools Virtual assistants (to provide reminders) 	<ul style="list-style-type: none"> AI to highlight interesting player interactions Data capture on player behaviour Player decision-making analysis 	<ul style="list-style-type: none"> Real-time language translation Collating past wargame data with open-source intelligence (OSINT) and secret information Speech transcription

Beyond specific AI applications, this report explores two possible investment pathways for AI in wargaming:

1. **Narrow, specialised AI applications for the near-term:** These are the simple AI use cases for near-term deployment. They include certain applications of domain-agnostic AI, for example real-time translation or automated transcription, as well as more specialised AI applications which can be deployed with minimal risk, for example procedural content generation for background content or natural language processing for data analysis.
2. **High-risk, high-reward AI investments:** While there were more polarised views on advanced AI applications such as decision aids and AI for adjudication, these warrant further exploration since they have high transformative potential. However, Defence should be cautious of investing in overambitious AI-based tools before fundamental technical and policy challenges have been addressed.

Several technical, policy, and ethical challenges were identified. For example, there is a need to ensure that the manner in which insights are gained from wargames can be

understood by all relevant stakeholders. AI could hinder this, making it harder for the inner workings of the wargame to be understood. Explainable AI that can demonstrate causality between decisions and impacts must therefore be prioritised. Additionally, AI may make it possible to quickly generate a game that appears compelling, but does not represent key concepts of value to the players and sponsors in a meaningful way. Efforts to identify how to build a trustworthy solution are necessary to enable adoption. Furthermore, AI that can demonstrate causality between decisions and impacts is required, as well as sovereignty discussions around the AI system and the access to the data underpinning the system. Issues around accountability and liability, as well as the potential de-skilling of senior decision-makers will also require consideration. There will be a need to psychologically prepare players for the integration of AI in wargaming. Finally, defence procurement and approvals delays would need to be overcome to be able to deploy an AI-enabled system in wargaming.

We conclude that the benefits AI can bring to wargaming could be significant, but there would be benefit in first introducing automation in specifically tactical or abductive wargames in the near term to manage risks. While some narrower applications of AI (e.g. speech transcription to support data analysts) can be deployed in the near-term, the most ambitious and transformative applications (e.g. Course of Action analysis and AI adjudication) require further research and investment. Similarly, further investment in cross-cutting enablers is required before AI can be introduced effectively into strategic-level wargames.

Recommendations

The recommendations summarised below focus on preparing the UK Defence community to develop and deploy effective AI-enabled wargaming systems in a responsible and evidence-based way:

1. **Invest in cross-cutting AI enablers** – The technical challenges identified in this study apply to various AI use cases, so cross-cutting enablers may be a more efficient use of resources to enable desirable applications. AI with contextual awareness, AI explainability techniques or AI that can establish causality between decisions and their effects would all help to unlock several AI use cases that could bring revolutionary benefits to decision-making.
2. **Develop verification and validation (V&V) procedures for AI wargaming tools** – Experts highlighted the prevalence of non-AI automation being marketed as ‘AI’, and risks around inexperienced customers procuring low-quality AI-enabled wargaming tools. Standardised V&V procedures will be important to ensure the reliability of AI tools for wargaming and their compatibility with rules of engagement, ethical standards, and explainability requirements.
3. **Trial narrow AI use cases in parallel to investing in complex AI-enabled wargaming** – There is a high cost to benefit ratio in the more sensationalised AI use cases such as CoA analysis and adjudication explored in this report. There is merit in testing implementation of narrow and relatively low risk automation and AI use cases (e.g. Procedural Content Generation, first pass scenario generation, and speech transcription) in wargames, while simultaneously investing in research to advance complex AI-enabled wargaming (e.g. AI for player decision-making analysis and AI in adjudication) in parallel.
4. **Commission research on wargaming epistemology and decision-making in wargaming** – A better understanding of what constitutes an effective wargame and successful human decision-making during a wargame would enable the design of AI-enabled tools that augment rather than add uncertainty to the validity of wargame outputs. In strategic wargames, a better understanding of quantifiable parameters that affect strategic outcomes would be required. The values that decision-makers apply in their decision-making would also need to be identified.

1. Introduction

Rapid technical advances have seen game AI make headlines globally, as AI agents have beaten human masters in ancient board games such as Go as well as strategic video games including Dota and Starcraft.¹ The Chinese company Tencent has also reportedly created AI agents that have beaten human players at Starcraft II.² Even complex features of human social interaction, including behaviours such as reciprocity and alliance formation, have been replicated by AI agents designed to play games like Diplomacy.³

In the UK Defence community, there is growing interest in applications of AI to wargaming to help address some of the novel challenges now facing strategic decision-makers, as well as some of the enduring challenges in implementing wargames.

Recent years have seen the complexity of the threat environment and the operational tempo increase rapidly, particularly in the context of sub-threshold and hybrid threats, as well as advances in emerging technologies. Consequently, UK Defence needs to grapple with overlapping, concurrent challenges, as well as novel escalation pathways within a context of uncertainty. Policymakers are now considering integrating AI in wargaming, both to prepare for the potential increased speed of decision-making required in AI-supported warfighting *and* in order to update the wargaming methodology itself. This report considers ways in which senior decision-makers can update their analytic toolbox using AI and automation to achieve decision advantage.⁴

There are some early signs of government, academic, and private sector exploration of AI to assist in the design, execution, and analysis of wargames as well as to support with wargaming logistics. Moreover, applications of AI in other sectors such as healthcare, diplomacy, education, and others reveal opportunities that may be leveraged in the wargaming context.

However, there is a lack of sufficient research evidence around fundamental questions including:

¹ James Goodman, Sebastian Risi, and Simon Lucas, "AI and Wargaming," *arXiv* (September 2020): 1-46.

² Peng Sun et al., "TStarBots: Defeating the Cheating Level Builtin AI in StarCraft II in the Full Game," *arXiv* (September 2018): 1-24.

³ Tom Eccles et al., "The imitation game: learned reciprocity in Markov games," *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (May 2019): 1934-1936.

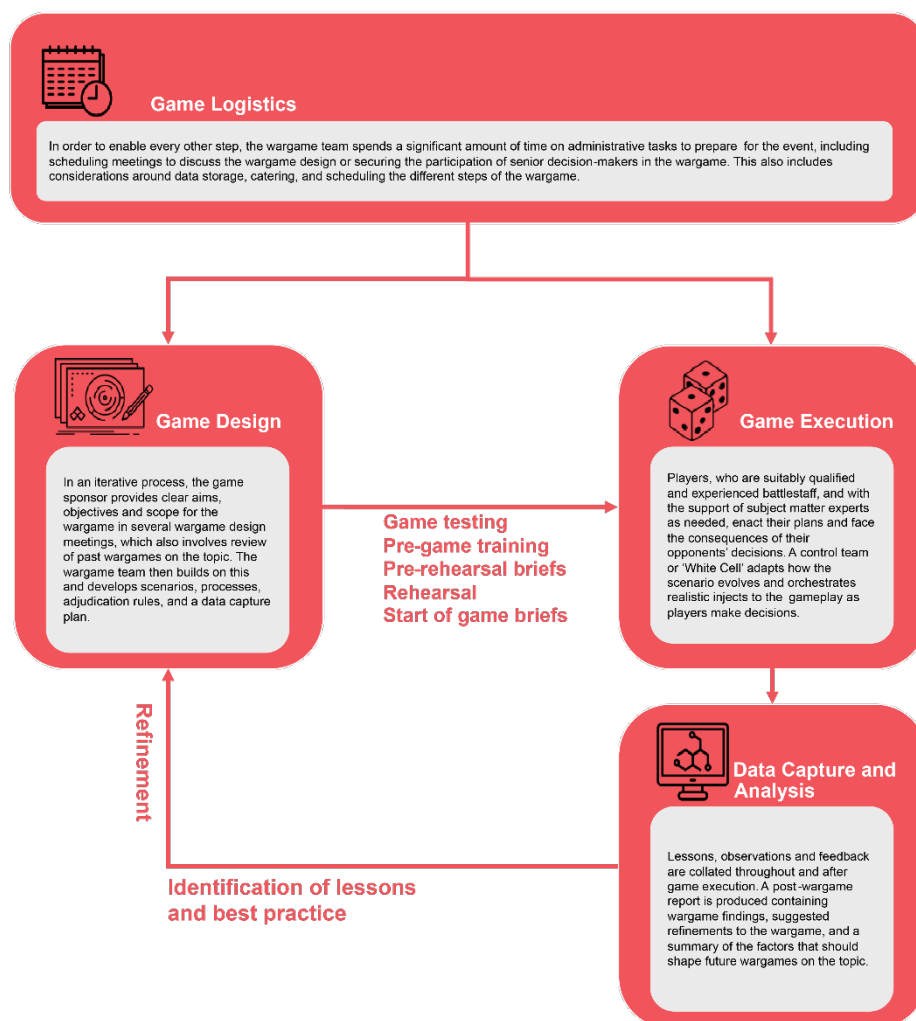
⁴ HM Government, *Defence Artificial Intelligence Strategy* (Ministry of Defence: 2022), <https://www.gov.uk/government/publications/defence-artificial-intelligence-strategy/defence-artificial-intelligence-strategy>.

- Whether AI really does support the outcomes of analytic games;
- Which wargaming contexts are most suited to AI deployment;
- What possible applications from non-defence sectors may be useful in segments of the analytic wargame workflow;
- Technical and policy challenges that need to be overcome to achieve the envisaged benefits that AI can bring to wargaming.

This study is intended to enhance understanding around the issues listed above.

The study considers the whole lifecycle of an analytic wargaming, from game logistics and game design, through to execution, data capture and analysis. Figure 1 below presents a simplified illustration of this analytic wargaming workflow.

Figure 1. Overview of wargaming workflow.



1.1. Research aims and methodology

This report aims to address the following research questions:

- **RQ1:** What is the current state-of-the art in AI-enabled systems that could support different stages in analytic wargaming?
- **RQ2:** Which non-defence decision-making contexts are leveraging AI-driven technologies and may be relevant comparators for analytic wargaming in the defence context?
- **RQ3:** What potential lessons and promising practice for the future of human-machine teaming in wargaming may be derived from:
 - International case studies of human-machine teaming in wargaming?
 - Relevant comparators to wargaming sub-components from other sectors and contexts?
- **RQ4:** Which AI applications for wargaming sub-components are showing the most promise?
- **RQ5:** What are the key challenges and obstacles to the introduction of AI to analytic wargaming workflows?

Data collection for this study was conducted over a four-month period from December 2022 – March 2023, including four core research activities:

1. **Literature review** covering technical, academic, and policy literature on topics such as AI in wargaming, game AI, policy and ethical challenges for wargaming and AI, and best practice in non-defence sectors.
2. **Semi-structured interviews** with 25 participants, including wargame designers, commercial game AI developers, cognitive psychologists, and modelling and simulation experts. Participants were identified using purposive, non-probabilistic sampling and interview data was analysed thematically using an inductive approach. Interviews were conducted on an anonymised, non-attributable basis.
3. **Case study analysis** exploring case studies that trialled the integration of AI in wargaming such as Group W's BrainSTORM, DARPA's Gamebreaker AI, and DARPA's Stabilizing Hostilities through Arbitration and Diplomatic Engagement

(SHADE) project. The study also examined non-defence applications of AI in commercial games, healthcare, and diplomacy, which may yield inspiration for segments of wargaming.

4. **Research workshop** attended by more than 40 representatives with expertise in professional wargaming, commercial game AI, modelling and simulation, non-defence AI, AI ethics and international law. The workshop focused on prioritising AI applications for wargaming into four categories: 'Invest', 'Deploy', 'Track', and 'Disregard'. Workshop data was analysed using a structured, inductive approach.

1.2. Caveats and limitations

In light of other ongoing work on this topic,⁵ the focus of this study was restricted to analytic manual wargames with some simulated elements (rather than fully automated wargames). The study conclusions therefore may not be generalisable to the methodology of wargaming as a whole. As emphasised throughout, the application of AI to wargaming is at an early, exploratory stage. This means that despite this report drawing on significant technical expertise, further work will be needed to scrutinise the applicability of the specific AI use cases discussed below to particular wargame contexts.

Critically, in writing this report, the research team has navigated significant variation in opinion among research participants. In particular, between professional wargamers on the one hand whose expertise in AI was frequently limited, modelling and simulation experts, and technical AI experts with no experience of wargaming. Participants had varied views on many fundamental topics from the definition of AI and how to categorise AI applications, to the potential utility of AI and other emerging technology across the wargaming workflow, yet there were several use cases that stakeholders from different backgrounds deemed to be worth tracking if not immediately deploying. A high-level analysis of the complete list of use cases in Appendix A informs the overall findings and recommendations on how to harness a diverse range of expertise to determine best practices for AI in wargaming.

⁵ Giles Ebbutt, "British Army aims for entirely new collective training capability," *Janes*, July 23, 2021, <https://www.janes.com/defence-news/news-detail/british-army-aims-for-entirely-new-collective-training-capability>; Lisa Baker, "Adoption of single synthetic environment predicted to save MoD £1.3bn, with total quantitative benefits reaching over £3bn," *Business in the news*, August 11, 2021, <https://businessinthenews.co.uk/2021/08/11/adoption-of-single-synthetic-environment-predicted-to-save-mod-1-3bn-with-total-quantitative-benefits-reaching-over-3bn/>; "Improbable's Defence business awarded contract for a second year," *Improbable*, December 11, 2020, <https://www.improbable.io/blog/improbable-uk-strategic-command-sse>.

1.3. Structure of this report

The remainder of this report is structured as follows. Section 2 will provide an overview of progress towards implementing AI-enabled wargaming, covering existing research and case studies, with a more comprehensive list of AI applications in wargaming provided in Appendix A. Section 3 will share findings from the research workshop and give detail on which applications of AI to wargaming are recommended for near-term and long-term prioritisation. Section 4 will cover the numerous technical challenges identified in the research, and Section 5 will synthesise the main policy challenges, before conclusions and recommendations are offered in Section 6. Appendix A contains the long-list of AI applications with a discussion of the envisaged benefits, challenges, and the rationale for their prioritisation or de-prioritisation in the defence wargaming context. Appendix B contains a glossary of the wargaming and AI terms used throughout this report.

2. Progress Towards Implementing AI-enabled Wargaming

Over recent years, interest in AI-enabled wargaming has ebbed and flowed. Now, there is an observable renewed interest in automating aspects of wargaming. This interest is largely driven by the emergence of AI applications in non-defence sectors – particularly the game AI sector, which could be applied to some segments of wargaming. This section contains findings and analysis on the state-of-the-art in introducing AI into wargaming based on a literature review, interviews, and challenge and validation through an interdisciplinary research workshop.

2.1. A renewed interest in automated wargaming

Interest in computer-assisted wargaming is by no means new. Interviewees described that during the 1950s-60s, ‘there was a sense that computers were growing up alongside gaming,’⁶ and as early as the 1980s, research was already investigating how to integrate wargaming with expert systems, a key branch of AI research at that time.⁷ This interest in automating wargaming was described by one interviewee as cyclical, rising and falling as decision-makers regularly become disillusioned with quantified approaches, but then seduced by the allure of novel technologies.⁸

In recent years, interest in computer-assisted wargaming has been revived. As one interviewee put it, the ‘current moment is analogous to the 80s, where there is a sense there have been technological advancements which we can return to.’⁹ Attention has now turned to utilising the latest available AI techniques – from reinforcement learning to generative AI – for defence decision-making.

Consensus on the current state of play was consistent: ‘AI for wargaming is in its infancy’,¹⁰ there is ‘not much going on’ in this space,¹¹ and ‘AI is not really used at this stage.’¹²

⁶ Author interview with think tank wargame design expert, 11 January 2023.

⁷ Paul K. Davis and Paul Bracken, *Artificial Intelligence for Wargaming and Modelling* (RAND Report, February 2022), https://www.rand.org/pubs/external_publications/EP68860.html.

⁸ Author interview (2) with think tank wargame design expert, 10 January 2023.

⁹ Author interview with think tank wargame design expert, 11 January 2023.

¹⁰ Author interview with private sector modelling and simulation expert, 1 February 2023.

¹¹ Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

¹² Author interview with think tank modelling and simulation experts, 2 February 2023.

However, there are now a few exploratory projects demonstrating nascent game AI or proposing potential AI applications for wargaming and simulation. Researchers have identified techniques such as deep learning as promising for future improvements to wargaming workflows,¹³ while others have pointed to using behaviour modelling for computer-generated forces,¹⁴ or to using AI techniques to identify which data from real-world exercises are most relevant to wargaming outcomes.¹⁵ This work is promising but nascent.

Despite the abundance of ambitious ideas, there remain significant doubts about whether any of these are 1) feasible or 2) helpful in answering decision-makers' questions. Many are dismissive of the potential for AI to resolve any of the pervasive challenges wargamers face.¹⁶ This scepticism came through strongly in interviews,¹⁷ which indicated that many wargamers did not support widespread AI deployment in wargaming (see Figure 2 below).

Some participants suggested that basic automation capabilities are frequently oversold as AI.¹⁸ Simultaneously, government customers' perceived lack of understanding of the inherent limitations of AI has contributed to confusion about which AI wargaming tools hold the most potential. For example, AI performs better at pattern detection than identifying causal relationships.¹⁹ AI performs better at quantitative analysis than qualitative, highly contextualised analysis.²⁰ Together, these two points of confusion – on the distinction between basic automation versus AI and on the sorts of tasks AI is suited to – often result in a debate which is insufficiently grounded in technical expertise and overly influenced by AI 'hype'.

At the same time, there were also optimists, and some of the AI applications discussed in Appendix A were already reportedly used by some wargame designers. For example, one interviewee discussed using ChatGPT to generate a first-pass version of background

¹³ James Goodman, Sebastian Risi, and Simon Lucas, "AI and Wargaming," *arXiv* (September 2020): 1-46.

¹⁴ Per-Idar Evensen et al., "Wargaming Evolved: Methodology and Best Practices for Simulation-Supported Wargaming," *Interservice/Industry Training, Simulation, and Education Conference* (December 2019): 1-13.

¹⁵ Cx_Wargaming, "Wargaming the Future (attribution AI and machine learning)," Video source, June 22, 2021, <https://www.youtube.com/watch?v=7DIPu8hAAU4>.

¹⁶ Jennifer McArdle and Eric Hilmer, "Effectively Integrating Technology into Wargames," *Interservice/Industry Training, Simulation, and Education Conference* (2022): 1-12.

¹⁷ Author interview with academic wargame design expert, 10 January 2023; Author interview (2) with think tank wargame design expert, 10 January 2023; Author interview with think tank wargame design expert, 11 January 2023; Author interview with former government wargame design expert, 18 January 2023; Author interview with think tank wargame design expert, 23 January 2023; Author interview with think tank wargame design expert, 24 January 2023.

¹⁸ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview (2) with think tank wargame design expert, 10 January 2023.

¹⁹ Author interview with think tank modelling and simulation experts, 2 February 2023.

²⁰ Author interview with government AI technical expert, 17 January 2023.

content for their games.²¹ Another interviewee discussed US-based efforts which claimed 'slow but good progress,' with one programme reportedly seeking to build a digital wargaming environment that combines a number of AI applications that several other interviewees said are not yet possible.²² This included AI for developing scenarios, for Course of Action (CoA) analysis, for red cell behaviour development (simulated adversaries) and AI to process data from wargames.²³ However, the same interviewee mentioned that these applications remain at the research stage and are not yet operational capabilities that could scale to the complexity of warfare.²⁴

²¹ Author interview with academic wargame design expert, 17 January 2023.

²² Author interview with government AI in wargaming expert, 26 January 2023.

²³ Author interview with government AI in wargaming expert, 26 January 2023.

²⁴ Author interview with government AI in wargaming expert, 26 January 2023.

Figure 2. Overview of headline messages on AI in wargaming from interview participants.

Professional wargame designers

AI doesn't fundamentally overcome challenges and designers need to be more honest about limitations of wargames.	It would be most useful to go through and do a list of pathologies where AI can go wrong.	AI is not magic, it can be a supplement to what you are doing but it will not be the cure-all.
Fix wargaming first.	If you have AI play both sides, it is no longer wargaming. It is M&S.	We don't have the foundations, but we are trying to build the cathedral anyway.
Don't take the people out of wargaming. Not saying no to putting AI in, but don't take the people out.	What you [Dstl] want is neither affordable nor feasible, so try plan B.	For senior decision-makers the real challenge will be balancing over-scepticism with over-optimism.

Modelling and simulation experts

It is really important to define the exact parameters of a game before you can think about AI.	Don't underestimate the difficulty of making AI work properly.	It doesn't make sense to integrate AI in a manual wargame - think digital.
With AI the buzz makes you forget that you are looking for an answer to one problem.	Think carefully about how adversaries are developing AI.	Dismiss AI at your peril. At the same time, overhype AI at your peril.

Non-defence AI experts

As a scientist I am a wishful optimist but many of these ideas would take serious development.	Make sure you are training AI to fight people not to fight a machine.	Explainable AI is the most important aspect as you need to know how AI outcomes are processed.
--	---	--

2.2. Case studies in AI-enabled wargaming

There are few case studies of AI-enabled wargaming in the public domain (with even fewer outside the confines of US government-funded research), and limited evidence on the

successes and failures of these projects. Details of the most prominent projects exploring AI-enabled wargaming are summarised below.²⁵ This gives an initial indication of how AI may be applied to wargaming, while also illustrating the exploratory nature of current projects. Notably, details of the methodology and challenges faced by these project teams are rarely published, making it difficult for others in the wargaming and modelling and simulation community to draw lessons from their work.²⁶

DARPA's Gamebreaker AI Exploration programme (May 2020 –?)²⁷

Project overview: Project using video games to study ways in which AI can be used within wargames.

Aim: To understand how AI might be used to tackle multiple problems in simulated games including AI to 1) assess game balance, 2) identify new tactics, and 3) test out what actions are most destabilising within video games. The ultimate aim is to use AI to create imbalanced games where one side is disadvantaged and so placed under significant stress. This is envisaged as desirable for training personnel.²⁸

Segment of wargaming: Game design and execution

Game mechanics: Nine teams were tasked with breaking two commercial video games, manipulating the games in order to identify unexpected states within games where game dynamics did not function as expected. The aim was to design a method that could extend beyond a single video game and possibly extend to wargaming scenarios. Different teams took different approaches. The Blue Waves team attempted to use autonomous game playing, interpretable AI and artificial neural networks to predict outcomes and determine why certain games were won or lost.²⁹ Meanwhile the Northrop Grumman team planned to model and break a complex simulator environment called 'Command: Modern Operations'.³⁰

²⁵ Other projects may also be of interest. These include projects developing synthetic wargaming environments (such as Improbable's Single Synthetic Environment or MITRE's JWARS project), projects exploring the impact AI will have on battlefield decision-making (for example, CSBA's Mosaic Warfare project, "exploiting AI and autonomous systems to implement decision-centric operations"), projects harnessing AI to improve decisionmakers' understandings of complex operational environments (for example, DARPA Causal Exploration) and other projects at the intersection of AI and wargaming which are yet to publish sufficient details in the project domain (for example, Project Voltron).

²⁶ Author interview with academic wargame design expert, 10 January 2023.

²⁷ "Gamebreaker AI Effort Gets Under Way," DARPA, May 13, 2020, <https://www.darpa.mil/news-events/2020-05-13>.

²⁸ "Gamebreaker AI Effort Gets Under Way," DARPA, May 13, 2020, <https://www.darpa.mil/news-events/2020-05-13>.

²⁹ "DARPA Gamebreaker AI Exploration Program," Blue Wave AI Labs, no date, <https://www.bluewaveailabs.com/darpa-gamebreaker/>.

³⁰ "Northrop Grumman Awarded DARPA Gamebreaker Contract," Northrop Grumman News Releases, August 12, 2020, <https://news.northropgrumman.com/news/releases/northrop-grumman-awarded-darpa-gamebreaker-contract>.

Group W's BrainSTORM (September 2018 –?)³¹

Project overview: A DARPA sponsored project based on STORM, the US Department of Defence model for multi-domain operations at the campaign level.³²

Aim of integrating AI: To assess whether AI can enhance course of action analysis and whether AI-assisted non-expert players can beat expert teams in a wargame.³³

Segment of wargaming: Game execution

Game mechanics: BrainSTORM consists of three distinct tools: 1) 'game cloud' is oriented towards facilitating high level strategy development, 2) 'color compare' facilitates the specification of these strategies to detailed CoA plans, 3) 'speedSTORM' is the environment in which detailed AI recommendations are accessed.

Evidence of effectiveness: Researchers reported that AI supported non-expert players can beat expert players. More tests are planned for the future.

DARPA's Stabilizing Hostilities through Arbitration and Diplomatic Engagement (SHADE) project (January 2022 – June 2023)³⁴

Project overview: This programme aims to use AI to improve strategic decision-making for diplomacy, in particular looking at negotiation processes. Specifically, a simulated environment of the classic game 'Diplomacy' is used to train and evaluate AI agents.

Aim: To investigate the extent to which AI can emulate human diplomatic behaviours including 'deception, collusion, profiling' and 'complex multi-party interactions'. Ultimately, the hope is this will improve understanding of the complex communication involved in human negotiations.³⁵

Segment of wargaming: Game execution

Game mechanics: This project is based on an open-source Diplomacy game engine which facilitates humans to play against Diplomacy bots. The team have created benchmark bots to assist in training new SHADE AI bots.

³¹ Alec Barker, "BrainSTORM: exploring artificially intelligent COA development in STORM," *SPIE.Library* (April 2021).

³² Multi-domain operations are those which span multiple domains, to include air, cyber/electromagnetic, maritime, land and space. Read more: HM Government, *Multi-Domain Integration* (Ministry of Defence: 2022), <https://www.gov.uk/guidance/multi-domain-integration>.

³³ Alec Barker, "BrainSTORM: exploring artificially intelligent COA development in STORM," *SPIE.Library* (April 2021).

³⁴ "Stabilizing Hostilities through Arbitration and Diplomatic Engagement," Project homepage, SHADE, no date, <https://www.shade-aie.org/>.

³⁵ There are significant parallels between this work taking place in the context of defence wargaming and work by DeepMind within game AI more broadly where researchers have studied the performance of AI agents in the board game of Diplomacy. See: Janos Kramar et al., "Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy," *Nature Communications* 13 (2022).

DARPA's Constructive Machine-learning Battles with Adversary Tactics (COMBAT) (May 2020 – ?)³⁶

Project overview: This programme aimed to develop AI applications which could model red team behaviour and be used to challenge blue teams in subsequent war games.

Aim: To develop several feasible AI-generated courses of action for the red team. And, subsequently to use these AI-based CoAs to determine optimum responses from the blue team. Ultimately, the aim is to use AI to help stimulate novel blue team behaviours.

Segment of wargaming: Game execution

Game mechanics: As part of this project, DARPA selected companies to develop AI applications and determine their own methodologies. Companies were tasked with developing their initial tactics according to “The Russian Way of War: Force Structure, Tactics and Modernization of Russian Ground Forces.”

DARPA's Strategic Chaos Engine for Planning, Tactics, Experimentation and Resiliency (SCEPTER) (January 2022 – ?)³⁷

Project overview: A project using automation and AI to explore the development of novel course of action plans at machine speed.

Aim: This project aimed to investigate whether the speed of CoA planning can be increased through advanced automation *and* whether these CoAs can compete with human plans.

Segment of wargaming: Game execution

Game mechanics: The strategies developed for machine-generated CoA planning will be tested within trusted simulated environments where they will be assessed by humans.

Insights from these defence case studies are summarised in Table 1 below. Some of the case studies were based within a fully digitised platform so the results were not indicative of possibilities for purely manual wargames or manual wargames with some digital elements. Moreover, there is limited evidence in the public domain of these projects' findings beyond

³⁶ “Constructive Machine-learning Battles with Adversary Tactics (COMBAT) (Archived),” DARPA, no date, <https://www.darpa.mil/program/constructive-machine-learning-battles-with-adversary-tactics>; John Keller, “Researchers eye embedding artificial intelligence into war games simulation to beef-up challenges,” *Military+Airspace Electronics*, June 3, 2020, <https://www.militaryairspace.com/computers/article/14177043/artificial-intelligence-ai-war-games-simulation>.

³⁷ John Keller, “DARPA SCEPTER project seeks to develop battle planning for complex military engagements at machine speed,” *Military+Airspace Electronics*, January 20, 2022, <https://www.militaryairspace.com/computers/article/14223907/battle-planning-complex-military-engagements-machine-speed>, <https://idstch.com/technology/ict/darpa-scepter/>; Managing Editor, “DARPA SCEPTER Program Will Explore Machine-Generated Strategies For Warfare,” *Defense Techconnect: The innovation source for national security*, January 24, 2022, <https://defensetechconnect.com/2022/01/24/darpa-scepter-program-will-explore-machine-generated-strategies-for-warfare/>.

self-reported claims on performance which are difficult to verify and to generalise. Nevertheless, some emerging lessons may be learnt.

Table 1. Lessons learnt from existing case studies of AI-enabled wargaming.

Project	Key findings
Gamebreaker AI	<ul style="list-style-type: none"> • Indicative of possibility of learning from commercial games industry successes, rather than starting from scratch for defence applications. • Highly quantified approach may be applicable to tactical games or games with a pre-determined rulebook. Unclear whether this is applicable to strategic wargames with high uncertainty.
BrainSTORM	<ul style="list-style-type: none"> • Early experiments show users, even non-experts, prefer to interpret AI recommendations themselves rather than adopting them wholesale.³⁸
SHADE	<ul style="list-style-type: none"> • Indicates efforts are being made to understand how AI agents might deal with human concerns at the more strategic end of the spectrum – for example, deceptive human behaviours. • Demonstrates the exploratory nature of much existing work, with the aim being to develop a prototype AI agent rather than a near-term capability.
DARPA COMBAT	<ul style="list-style-type: none"> • Illustrates benefits of competitive research models to drive innovative ideas within private sector. • Unclear if this project demonstrates a successful application of AI red team agents given the lack of public evidence of which AI methodologies were applied or what results were achieved.
DARPA SCEPTER	<ul style="list-style-type: none"> • This project is at an early stage and further developments should be tracked for further insights.

2.3. Opportunities in non-defence applications of AI

This study also investigated where wargaming may take inspiration from non-defence sectors. While non-defence AI use cases may appear alluring, it is necessary to bear in mind three key caveats before drawing conclusions based on these applications:

³⁸ Alec Barker, "BrainSTORM: exploring artificially intelligent COA development in STORM," *SPIE Library* (April 2021).

1. The defence context comes with specific requirements of AI assurance given the grave consequences of any errors in decision-making,³⁹ and given the number of novel situations which may be encountered.⁴⁰
2. Wargames are highly complex and human centric. This complexity is often not reflected in other sectors (even in complex games such as Diplomacy).⁴¹
3. Budgetary constraints differ enormously by sector. The cost to reward ratio required to make defence investment worthwhile will differ from other commercial sectors.⁴²

Nevertheless, certain lessons can be learnt from other sectors, both in cases where AI applications from elsewhere may be adapted for Defence *and* in cases where challenges faced elsewhere are indicative of similar barriers faced by Defence.

When asked which sectors would be most relevant to advancing AI-enabled wargaming, interview participants mentioned a wide range of sectors (as illustrated in Figure 3 below).

Figure 3. Most frequently identified comparator sectors based upon interview data, illustrating that 'healthcare' and the 'commercial games industry' were most frequently cited by interviewees as relevant to defence applications.



³⁹ Author interview with private sector modelling and simulation expert, 15 February 2023.

⁴⁰ Author interview with government wargaming experts, 14 December 2023.

⁴¹ Author interview with think tank wargame design expert, 23 January 2023; Author interview with private sector modelling and simulation expert, 1 February 2023.

⁴² Author interview (2) with AI technical expert, 11 January 2023.

In addition to these varied sectors identified in interviews, further relevant sectors for Defence to track have been identified through literature review. These include:

- Applications of AI in insurance, where progress has been made towards using chatbots to provide targeted information to customers,⁴³
- Commercial risk management, where AI has helped to develop dashboards which summarise up-to-date global events,⁴⁴
- Applications in the education sector where algorithms used to grade students may provide insights for those wishing to use AI in player decision-making analysis.⁴⁵

All these AI use cases and more have influenced and inspired the applications of AI to wargaming discussed in Appendix A. However, three sectors were particularly prominent and the relevance of each to AI-enabled wargaming is summarised below. For each sector we indicate significant AI use cases drawn from relevant academic literature, while also pointing to their specific relevance to AI-enabled wargaming.⁴⁶

Table 2. Relevant AI applications from the healthcare sector.

AI use cases	Relevance to AI-enabled wargaming
AI to assist clinicians with triage decisions in emergency care ⁴⁷	AI for decision support in healthcare may be indicative of technical possibilities for similar applications to wargaming, in particular promising AI use cases to support triage decisions by predicting the likelihood of severe deteriorations in patient condition. ⁴⁸ Similar techniques may be applied to course-of-action analysis, as this is also an environment where AI assists humans in choosing which actions are favourable in a high-stakes context based on predictions of possible associated risks.

⁴³ Benno Keller, "Promoting responsible artificial intelligence in insurance," *The Geneva Association* (April 2020).

⁴⁴ Fahim K Sufi, "AI-GlobalEvents: A Software for analyzing, identifying and explaining global events with Artificial Intelligence," *Software Impacts* 11 (February 2022): 1-5.

⁴⁵ Victor Gonzalez-Calatayud, Paz Prendes-Espinosa and Rosabel Roig-Vila, "Artificial Intelligence for Student Assessment: A Systematic Review," *Applied Sciences* 11, no. 12 (June 2021): 5476.

⁴⁶ Given the breadth of sectors addressed here, AI use cases will necessarily be discussed in broad terms and further work will be needed to explore how lessons can be learnt from specific projects from outside of Defence.

⁴⁷ Marta Fernandes et al., "Clinical Decision Support Systems for Triage in the Emergency Department using Intelligent Systems: a Review," *Artificial intelligence in medicine* (November 2019); Aidan McParland and Kiran Grant, "Applications of artificial intelligence in emergency medicine," *University of Toronto medical journal* 96, no. 1 (January 2019).

⁴⁸ Aidan McParland and Kiran Grant, "Applications of artificial intelligence in emergency medicine," *University of Toronto medical journal* 96, no. 1 (January 2019).

AI chatbots in psychology and decision support tools to guide psychologists or psychiatrists during therapy sessions ⁴⁹	Research on AI in psychology points to numerous ethical and policy challenges associated with decision support tools, from a persistent lack of trust to cognitive de-skilling of practitioners. The defence sector can draw on this work for virtual assistants that provide legal reminders or CoA.
--	---

Table 3. Relevant AI applications from the commercial game sector.

AI use cases	Relevance to AI-enabled wargaming
Deep learning for procedural content generation (PCG) ⁵⁰	Progress towards using deep learning for PCG is significant to wargame-design as the sophistication of AI-designed game elements, such as non-player characters, is likely to increase.
AI game agents, including for games such as Dota, Starcraft, and Go ⁵¹	AI game agents should be tracked by those interested in using AI to automate red or blue teams in wargaming. However, the decision space of these games is typically narrower, meaning lessons are not directly transferable. Additionally, the end goal of this research is often simply to create human-beating AI agents, whereas in wargaming, greater emphasis is placed on the insights gained by analysing human actions within the game itself.
AI for the boardgame 'Diplomacy' ⁵²	Certain complex games show a higher relevance to the context of wargaming, with DeepMind research on Diplomacy showing particular promise given progress towards emulating human behaviours such as negotiation and alliance formation. ⁵³ Results show promise with regard to AI agents' abilities to act both cooperatively, forming alliances, and duplicitously, breaking alliances.

⁴⁹ Christopher Burr and Rosamund Powell, "Trustworthy Assurance of Digital Mental Healthcare," *The Alan Turing Institute* (November 2022).

⁵⁰ Jialin Liu et al., "Deep Learning for Procedural Content Generation," *arXiv* (October 2020).

⁵¹ David Silver et al., "Mastering the game of Go with deep neural networks and tree search," *Nature* 529 (January 2016): 484-489; Christopher Berner et al., "Dota 2 with Large Scale Deep Reinforcement Learning," *ArXiv* (December 2019); Dan Garisto, "Google AI beats top human players at strategy game *StarCraftII*," *Nature News* (October 2019); Janos Kramar et al., "Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy," *Nature communications* 13 (December 2022): 7214.

⁵² Janos Kramar et al., "Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy," *Nature communications* 13 (December 2022): 7214.

⁵³ Janos Kramar et al., "Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy," *Nature communications* 13 (December 2022): 7214.

Table 4. Relevant AI applications from diplomacy and conflict prevention.

AI use cases	Relevance to AI-enabled wargaming
Research has explored AI for decision support in the context of diplomacy, for instance using AI to provide recommendations to ministries on strategic decisions. ⁵⁴	Work on AI decision support for diplomacy will be of particular relevance to those wishing to use AI to provide virtual assistants for wargame players, which consider complex international political environments. For example, tools to remind players of the legal context relevant to a particular game.
AI to analyse transcripts during diplomatic negotiation. ⁵⁵	AI for analysing transcripts in the context of armed conflict prevention is indicative of the potential of AI to assist in analysis of wargaming transcripts, particularly to analyse convergence and divergence in viewpoints between teams over the course of a wargame.

This overview of progress towards AI-enabled wargaming illustrated that while significant progress has been made in adjacent sectors, much of this is as yet untapped within wargaming itself. As the remaining sections will reiterate, further research is needed to assess the extent to which existing AI techniques may be practical within wargaming workflows *and* the extent to which they can help to address some of the persistent challenges of analytic wargaming.

⁵⁴ Corneliu Bjola, "Diplomacy in the Age of Artificial Intelligence," *EDA Working Paper* (January 2020).

⁵⁵ Author interview with AI technical expert, 10 January 2023; Miguel Arana-Catania, Feliz-Anselm van Lier and Rob Procter, "Machine Learning for Mediation in Armed Conflict," *arXiv* (August 2021).

3. Promising Applications of AI in Wargaming

Despite broad variance of opinion on the utility of AI in wargaming, and a lack of consensus on which AI applications hold the most utility, even the more cynical interview participants identified particular contexts in which AI may bring benefits to wargaming. Figure 4 below presents the longlist of AI use cases identified, which are discussed in further detail in Appendix A. This section only focuses on the list of use cases prioritised in a March 2023 workshop.

Figure 4. Long-list of AI use cases for wargaming identified

Game design	Game execution	Game analysis	Game logistics
AI red team agents	Simulating opponents' reactions	AI to highlight interesting interactions between teams	Real-time language translation
Procedural Content Generation	Adjudication	Data capture on player behaviour	Collating past wargame data with OSINT and secret information
Scenario and inject generation	AI to run through tactical decisions	Player decision-making analysis	Speech transcription
Gameplay testing	AI decision assistants and CoA analysis tools		
	Virtual assistant (to provide legal reminders)		

AI was assessed to be useful in narrow and more specialised applications like tactical gameplay and abductive games, as well as AI applications for supporting data analysis – where the main risks to the quality of gameplay can be managed. Abductive games are games with a known end state where the objective of the game is to identify the best possible explanation for a surprising fact or observed phenomenon. Tactical and abductive games lend themselves more to the integration of AI because the possible outcomes and impacts of decisions are more often known and quantifiable. There were also some high-risk, high-reward applications mentioned. These would require significant further investment and research, but some study participants believed they could contribute to breakthroughs in decision advantage through wargaming.

3.1. AI in narrow and specialised applications

The most readily deployable AI solutions do not alter integral game mechanics, but could instead improve the logistics of the game, support data analysis, or add more realistic texture and noise to the game (e.g., synthetic Tweets or diplomatic cables that add tension to the game). Speeding up processes like automated speech transcription and language

translation during a wargame using automation, or using generative AI for a first draft of scenarios was considered by interviewees to present relatively low risk.⁵⁶ Often, these more practical applications were not at the cutting edge of AI research, and several interviewees cited simpler non-AI automation techniques as most useful.⁵⁷ However, while simple automation could provide significant efficiency gains, a preoccupation with falsely marketing these automation techniques as AI has contributed to hype and confusion.⁵⁸ Greater technical literacy is needed to distinguish between AI and simpler automation techniques, and to identify which option is best applied to a particular wargaming context.

*'We need more mini models rather than large command rooms with complex simulations. This could bring benefits of modelling to a great variety of roles in wargaming.'*⁵⁹

One of the reasons for the limited progress in integrating AI into wargaming is that ambitions for these systems are too high, with visions of automating all aspects of wargaming from game preparation through to execution and analysis. Instead, AI development for wargaming should be targeted for specific purposes.⁶⁰

For example, in data capture and analysis for analytic wargames, interviewees highlighted opportunities to reduce analysis time by increasing the efficiency of data capture on what transpired in the game, or collating wargame data, secret information and open-source intelligence.

Introducing AI to take over tactical decisions could also potentially reduce the number of personnel required in a wargame. In a strategy game where the effect is distant from the cause, and complex battlespaces are orders of magnitude larger, it would be more difficult to envision reliable AI. AI is best suited to making relatively simple decisions such as firing policy, where cause and effect are seconds apart and the simulation can be run iteratively.

Moreover, the more abductive, quantifiable, and rules based the game is, the more likely AI might be able to help in a simulated environment.⁶¹ The more strategic the game, the fuzzier

⁵⁶ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview with think tank wargame design expert, 10 January 2023.

⁵⁷ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview with think tank modelling and simulation expert, 17 January 2023; Author interview with private sector modelling and simulation expert, 1 February 2023.

⁵⁸ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview with think tank modelling and simulation expert, 17 January 2023; Author interview with private sector modelling and simulation expert, 1 February 2023.

⁵⁹ Author interview with government AI technical expert, 19 January 2023.

⁶⁰ Author interview with think tank wargame design expert, 24 January 2023.

⁶¹ Author interview with private sector modelling and simulation expert, 1 February 2023.

the causality becomes, leaving many aspects of the game open to interpretation and unknown outcomes. As one interviewee stated, 'If AI can solve it for you, you don't need a wargame, you could just run a simulation.'⁶²

This illustrates the importance of clarifying the aim of a wargame before considering the introduction of AI. If a wargame is being used to understand other human adversaries' decision-making, AI resolving the problem for the player will not bring the same sorts of insights as humans tackling the problem themselves. However, if the game is seeking to understand how to compete with human adversaries who are assisted by AI-enabled CoA analysis, then some of the more complex AI use cases in Section 3.2 may be worth exploring.

3.2. High-risk, high-reward AI applications in game execution and data analysis

There were a few bleeding-edge AI applications that interviewees believed could constitute revolutionary innovation to the type of data that can be extracted from a wargame. These include decision aids that support course-of-action (CoA) analysis – potentially through virtual assistants and chatbots, AI in adjudication, and AI that enables analysis of player decision-making through a combination of computer vision and Natural Language Processing (NLP). These revolutionary applications could be bending the possible definitions of a wargame itself and the potential benefits could be significant.

However, these potentially high reward applications were also the most disputed, and some warned that developing AI tools for these applications could be the most expensive and least effective way to pursue decision advantage.⁶³ Some described this challenge as intractable, especially given the sub-optimal results emerging from US projects, despite such significant investments.⁶⁴ However, though this innovation approach may necessitate a long road of failures before these use cases are successfully demonstrated, investment in these use cases could also generate revolutionary advances that constitute a competitive edge in strategic competition.

⁶² Author interview with think tank wargame design expert, 24 January 2023.

⁶³ Author interview with think tank wargame design expert, 23 January 2023.

⁶⁴ Author interview with think tank wargame design expert, 11 January 2023.

AI for CoA analysis and other decision aids could help speed up decision-making, or help players consider novel tactical options and present their likely results.⁶⁵ There are some emerging projects such as DARPA's Strategic Chaos Engine for Planning, Tactics, Experimentation and Resiliency (SCEPTER) project, which seeks to explore the use of decision aids producing machine-generated strategies to compete with humans in simulated environments, but little is known of the concrete results.⁶⁶

In order to calibrate the right level of trust in AI-enabled CoA, there is a need for players, game sponsors and overseers to understand how AI has arrived at its outputs, but the state-of-the-art in AI is insufficiently transparent to enable this. Players and data scientists seeking to understand the games also need to understand why they succeeded or failed.⁶⁷ Moreover, the data truly needed to bring value to a wargame includes insights such as which CoAs were discarded and for what reason. This may be challenging for AI-enabled CoA, which may not be able to fully explain its reasoning.⁶⁸

Adjudication is one arena in which expert viewpoints diverged significantly, with some suggesting machine learning might be effectively leveraged to determine outcomes in novel situations, while others warned strongly against this. Adjudication is usually done by developing a results table of known outcomes and rolling the dice, and there is interest in integrating AI to make adjudication less random and more predictive based on historical analysis.⁶⁹ But this only works well if the potential consequences of players' actions are well known and understood.⁷⁰ If the game in question is an analytic game aimed at understanding the potential decision-making of an adversary in a novel situation, then everyone is equally inexpert, and by definition, there will not be a results table that can be programmed into the AI.⁷¹ Moreover, computerised adjudication is relatively inflexible if new insight is revealed throughout the wargame, requiring redesign of the scenario or rules of the game.⁷² In manual wargames, humans with expert knowledge can quickly adapt the rules, scenario and the next script of events to the decisions of players. This does not mean that human adjudication is perfect, but instead that any plans to integrate AI will need to

⁶⁵ Author interview with academic wargame design expert, 30 January 2023; Author interview with government AI in wargaming expert, 26 January 2023.

⁶⁶ John Keller, "DARPA SCEPTER project seeks to develop battle planning for complex military engagements at machine speed," *Military + Aerospace Electronics*, January 20, 2022, <https://www.militaryaerospace.com/computers/article/14223907/battle-planning-complex-military-engagements-machine-speed>.

⁶⁷ Author interview with think tank wargame design expert, 24 January 2023.

⁶⁸ Author interview with think tank wargame design expert, 24 January 2023.

⁶⁹ Author interview with government AI technical expert, 19 January 2023.

⁷⁰ Author interview with academic wargame design expert, 30 January 2023.

⁷¹ Author interview with academic wargame design expert, 30 January 2023.

⁷² Author interview with think tank wargame design expert, 24 January 2023.

ensure the sort of continued flexibility in adjudication outcomes which can currently only be achieved with expert involvement.

Another challenge is that for both CoA analysis and adjudication, there are innumerable possible environmental factors, and no existing datasets to train AI on all possible outcomes by several potentially illogically behaving agents.

4. Technical Challenges and Requirements

This section examines the key technical challenges and requirements for effectively integrating AI into wargaming based on inputs from the literature review, interviews and workshop. These include the intricacy of developing a valid methodology, risk of users over-trusting AI system outputs, ensuring trustworthiness, establishing causality, providing AI systems with an appropriate degree of contextual awareness and considerations around AI system and data sovereignty. These challenges and requirements are each considered in turn.

4.1. Development of a valid methodology

Constructing a wargaming methodology that yields logically valid outputs is an intricate art. Introducing AI in a way that does not further add doubt to the validity of the outputs of the analytic wargame will be crucial to the adoption of AI in wargaming. Humans are not rational decision-makers, and when humans are involved, game theory rarely plays out in practice and results are always different. This led a wargame designer to suggest it would be difficult for AI to replicate human decision-making patterns in AI red team agents or in predictive analytics,⁷³ particularly without labelled datasets on adversary decision-making patterns.⁷⁴ While the emergent behaviour of AI may appear complex, it may be derived from only basic rules, whereas human behaviour can be creative and unpredictable.⁷⁵ This perceived inability to mimic the emergent behaviour of humans is seen as a disadvantage by many when the purpose of the wargame is to improve capabilities against humans, not machines.⁷⁶

Some envisaged that reinforcement learning may mitigate this need for data or help enumerate every possibility in ways humans cannot, and that AI creativity could advance in ways that overcome these limitations. It should be noted, however, that research in this area is nascent.

In AI techniques like agent-based modelling, it is important to model features of the environment and if all the relevant factors are not modelled, then the agent behaves in ways that do not make sense and therefore make the results of the wargame invalid.⁷⁷ In manual

⁷³ Author interview with government wargaming experts, 10 December 2022.

⁷⁴ Author interview with private sector modelling and simulation expert, 15 February 2023.

⁷⁵ Author interview with private sector modelling and simulation expert, 15 February 2023.

⁷⁶ Author interview with government AI technical expert, 17 January 2023.

⁷⁷ Author interview with private sector modelling and simulation expert, 15 February 2023.

wargames, it is possible to have human players replicating the decisions of hostile states, but even experts will disagree on whether these types of wargames reflect the ground truth. One interviewee emphasised:

‘Wargaming has poorly defined error bars. It’s inaccurate because it’s making some predictions about scenarios and the end point is never going to happen because there are so many decisions that would need to be made in the real world.’⁷⁸

Some interviewees highlighted that there may be a temptation for commercial wargame suppliers who lack the background in wargame design to rush to develop the code for AI in wargaming before the validity of the methodology is established.⁷⁹ In order to overcome this challenge, more research is required to understand how theoretical models behind wargames can be integrated with AI and still be intellectually defensible. This means setting definitions that can be coded for nebulous concepts like ‘democracy’⁸⁰ and quantitatively optimising for these definitions. At the same time, there is a potential trade-off between optimal, yet brittle and predictable solutions with set definitions and more robust solutions that allow the AI to explore without being constrained by narrow definitions. The key will be knowing the right contexts to prioritise optimal or robust solutions.

These risks could be further exacerbated by government customers’ limited technical understanding of AI. One interviewee with experience trialling the integration of AI in a simulation said he ‘felt like the wizard of Oz behind the curtain,’ because what players in the wargame assumed the computer did was far more complex and sophisticated than what the computer actually did.⁸¹

‘We don’t have the foundations, but we are trying to build the cathedral anyway.’⁸²

4.2. Over-trust

The epistemological validity of the outputs of a wargame and the capacity of different types of AI may be variably understood across UK Defence, especially at the senior levels.⁸³ When players are immersed, they may stop critically assessing the model and place too much trust

⁷⁸ Author interview with think tank wargame design expert, 24 January 2023.

⁷⁹ Author interview with think tank wargame design expert, 23 January 2023.

⁸⁰ Author interview with think tank wargame design expert, 23 January 2023.

⁸¹ Author interview with private sector modelling and simulation expert, 15 February 2023.

⁸² Author interview with think tank wargame design expert, 23 January 2023.

⁸³ Author interview with think tank wargame design expert, 10 January 2023; Author interview with think tank wargame design expert, 23 January 2023.

in the outputs.⁸⁴ An illustrative example is in AI-enabled red team agents, which simulate the actions of opponents, but which are difficult to validate for realism.⁸⁵ Another example is in generative AI for scenario generation, which may not be basing its outputs on a sufficiently credible or robust evidence base.⁸⁶ Immersed players may forget to challenge the validity of AI outputs. Wargames are not meant to be exact predictions and traditionally adjudicated wargames are not always challenged either, but AI may add further uncertainty to the defensibility of wargame outputs.

In order to develop a wargame with valid outputs, the theoretical model underpinning the game needs to be clear to policymakers and developers.⁸⁷ There is therefore a need to cultivate an awareness that depending on different epistemological positions, the resulting benefits of a wargame or simulation are considered differently.

4.3. Ensuring trustworthiness

Even if the technical barriers to some of the AI techniques discussed in the previous section were to be overcome, there would remain several ethical, legal and political challenges which would need to be overcome in order to engender widespread trust in decisions made on the basis of AI-enabled wargames. Some interviewees believe that integrating AI in wargaming is far enough from high-risk decisions to be relatively low risk,⁸⁸ but others contended that there are still ethical risks around taking conclusions from an oversimplified synthetic decision-making context and applying negative learning in the operational environment outside the wargame.⁸⁹ On the other hand, it would be near impossible to determine whether it was the wargame or other factors that affected a senior decision-maker's decision-making.

Furthermore, it is unclear how ethics can be encoded in some use cases suggested, such as virtual assistants that provide reminders on ethical considerations, since sufficiently ethical actions in every given context are subjective and not easily quantifiable thresholds that can be coded into a machine.⁹⁰

⁸⁴ Author interview with government AI technical expert, 19 January 2023.

⁸⁵ Author interview with academic wargame design expert, 18 January 2023, Author interview with government AI technical expert, 17 January 2023.

⁸⁶ Author interview (2) with AI technical expert, 11 January 2023.

⁸⁷ Author interview with think tank wargame design expert, 23 January 2023.

⁸⁸ Author interview with government wargaming experts, 10 December 2022.

⁸⁹ Author interview with think tank wargame design expert, 11 January 2023; Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

⁹⁰ Author interview with government AI technical expert, 19 January 2023.

Moreover, some stakeholders anticipated difficulty in securing senior decision-makers' buy-in for AI-enabled systems that could invasively collect and analyse data on their decision-making performance, particularly because simulated models have shown low reliability to date.⁹¹

4.4. Establishing causality

The main issue that makes it difficult to transform strategic level wargaming using AI applications is the difficulty of predicting the causal chain of events between decisions and outcomes at the strategic level. Many of the AI techniques raised by interviewees (see Appendix A for a comprehensive list), from natural language processing to reinforcement learning, are more proficient at finding correlations,⁹² and fuzzier at establishing the causal effects of decisions.⁹³ Developing AI that can understand and apply the rules of the strategic context to predict these causal chains at the complexity of strategic level decisions, including escalatory and de-escalatory moves, would be a game-changing capability.⁹⁴

4.5. AI with contextual awareness

In order to enable human-machine teaming in AI use cases like CoA analysis and adjudication, the human and the AI tool would need to have the same situational awareness about what is happening on the manual boardgame if the whole system is not digitised.⁹⁵ Moreover, to realise the envisaged benefits of use cases like AI highlighting interesting interactions between red and blue human agents to data scientists,⁹⁶ AI would need to be able to recognise the most salient interactions of the players. There is ongoing, nascent research on informing AI about the context of moves in a wargame, but this is currently not yet possible in all contexts.⁹⁷ There will be a need to invest in human-machine interfaces that enable players and their deliberations and actions to be integrated into the digital space where AI can process it in a fully simulated environment.

⁹¹ Author interview with private sector modelling and simulation expert, 15 February 2023.

⁹² See Appendix B for further details on the AI techniques seen as most relevant to AI-enabled wargaming.

⁹³ Author interview with think tank modelling and simulation experts, 2 February 2023.

⁹⁴ CETaS Workshop, 16 March 2023.

⁹⁵ CETaS Workshop, 16 March 2023.

⁹⁶ Author interview with government wargaming experts, 10 December 2022; Author interview (2) with think tank wargame design expert, 10 January 2023.

⁹⁷ Author interview with private sector modelling and simulation expert, 1 February 2023.

4.6. Data and AI system sovereignty

If the skilled talent to develop AI-enabled wargaming systems exists in the commercial sector or in other nations, then the data and tools underpinning the systems could enable other nations and non-government actors to access sensitive information.⁹⁸ This may run contrary to the pursuit of strategic autonomy on AI for strategic advantage and could generate security risks. However, in order to build sovereign AI capabilities, there may be a need to collate large volumes of data across national boundaries, which some nations are addressing by coordinating at a European level. The UK will need to consider its own approach to data and AI sovereignty in relation to military AI and wargaming specifically.

⁹⁸ Author interview with think tank modelling and simulation experts, 2 February 2023.

5. Policy and Ethical Challenges

This section discusses the policy and ethical challenges identified in the research which must be overcome in order to implement effective AI tools in wargaming. The analysis in this section is based on inputs from the literature review, interviews and interdisciplinary workshop.

5.1. Accountability and liability

The first policy challenge identified in the research relates to accountability of decision-making. Decisions made in defence contexts may have life-or-death consequences. At the extreme end of the spectrum, if it is impossible to determine whether AI's integration in CoA analysis affects high risk decision-making, there is a risk that the decision-making informed by AI could cause loss of life, but that the AI's influence on this outcome cannot be proven or challenged. Such CoA systems should therefore not be deployed until the causality between the experience of AI in wargaming and its effects on subsequent decision-making are better understood.

5.2. Reduction of decision-making pressure on players

Another policy concern is whether introducing AI in wargaming or in other decision aids could have a detrimental effect on senior decisionmakers' decision-making performance due to the reduction of pressure in the game environment.⁹⁹ The introduction of AI use cases like virtual assistants that provide reminders on legal and regulatory frameworks, historical data or past wargames may reduce the risk of cognitive overload,¹⁰⁰ but may also create cognitive dependency on the tools. At the same time, one interviewee contended that there is such a shortage of personnel in Defence that the risk of *not* trying to leverage AI is more pressing than the risk of cognitive deskilling.¹⁰¹

⁹⁹ Author interview with academic wargame design expert, 10 January 2023.

¹⁰⁰ CETaS Workshop, 16 March 2023.

¹⁰¹ Author interview (2) with think tank wargame design expert, 10 January 2023; Author interview with think tank wargame design expert, 11 January 2023; Author interview with academic wargame design expert, 17 January 2023.

5.3. Psychologically preparing players for human-machine teaming in wargaming

One possible way of helping humans build calibrated trust in a human-machine teaming system in wargaming is by trialling AI applications in narrow applications to increase players' experience with AI. This process could begin with automating simple manual processes, before gradually introducing more complex AI.¹⁰²

*'Experience is the antidote to mysticism. The more the services get their hands on and experiment with these tools, the less they will be this mythical thing on a pedestal.'*¹⁰³

As AI is integrated in hybrid and partially simulated manual wargames, there will be a need to determine the appropriate roles for humans and AI. The human's role will likely be more centred on situational awareness, observing how well AI is undertaking its tasks. Introducing AI in innocuous use cases like real-time language translation in multi-national wargames, then gradually introducing more advanced AI applications like AI for analysis of player decision-making patterns could help facilitate this transition.

5.4. Efficient defence procurement and approvals processes

An AI-based system for wargaming will need to be approved through slow bureaucratic approvals processes and be able to navigate approvals for access to data from classified and international networks,¹⁰⁴ which will all generate delays. For example, interviewees discussed the utility of using AI to collate past wargame data with secret intelligence and OSINT to automate risk analysis (e.g., via artificial neural networks),¹⁰⁵ but approvals for all these types of data will likely create delays to model development.

'The problem you run into is for it to be genuinely labour saving it needs to be intuitive enough and work within the hellscape of any defence department's IT system. By the time you are navigating classified networks, international networks,

¹⁰² Author interview with think tank wargame design expert, 23 January 2023.

¹⁰³ Author interview with private sector modelling and simulation expert, 15 February 2023.

¹⁰⁴ Author interview with think tank wargame design expert, 24 January 2023.

¹⁰⁵ Author interview with government AI technical expert, 17 January 2023.

*different ministries, and departments... [there is a possibility that] a tool will be more trouble than it is worth.*¹⁰⁶

This challenge is not unique to AI, and there are ongoing efforts to address these procurement challenges across UK Defence. But if these delays are not addressed, even an intellectually defensible and technically feasible AI system may not be delivered at the point of need – weakening the UK’s ability to secure strategic advantage over its adversaries.

Finally, appropriate upskilling of procurement professionals will be required to ensure those procuring AI systems for defence can meaningfully scrutinise the system’s capabilities and limitations. Professional wargamers suggested that often, one of the largest causes for delay was in government customers ‘asking the wrong questions’ at the outset of wargaming projects, and AI is unlikely to help them identify the right questions and objectives for wargames faster.¹⁰⁷

¹⁰⁶ Author interview with think tank wargame design expert, 11 January 2023.

¹⁰⁷ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview with think tank wargame design expert, 10 January 2023; Author interview with academic wargame design expert, 30 January 2023.

6. Conclusions and Recommendations

There are numerous ways in which AI may be integrated into analytic manual wargaming, but these use cases vary significantly in their technical complexity, practical feasibility, and level of operational risk. As such, this report has sought to distinguish between unrealistic use cases, and more realistic use cases where potential risks could be best managed, and where the benefits to UK decision superiority are significant enough to warrant further exploration.

The AI use cases discussed throughout this study (which are discussed along with their specific benefits and challenges in Appendix A) reflect a potential revolutionary change in wargaming, which is pointing to hybrid wargames and a potential bridge between the wargaming and modelling and simulation communities, which have historically come from competing epistemological views on the value of wargaming. It is perhaps in this breaking down of barriers between communities that senior decisionmakers will be able to inject more ingenuity of thinking and speed into decision-making, to contribute towards operational advantage.

In conclusion, this study found that AI has potential as a *tool* to supplement some low-risk elements of wargaming, but introducing AI won't overcome the delays and costs often associated with wargaming. Introducing AI in wargaming itself may in fact introduce *new* costs, administrative delays, and uncertainty in outputs. However, AI may still contribute towards decision *advantage* by enabling new methods of data analytics for decisionmakers and enabling analysis of player decision-making, if some of the enabling AI techniques discussed in this report are advanced to become operationally viable.

The following recommendations are focused on preparing the UK Defence community to develop and deploy effective AI-enabled wargaming systems in a responsible and evidence-based way.

Recommendations

1. **Invest in cross-cutting AI enablers** – The technical challenges identified in this study apply to various AI use cases, so cross-cutting enablers may be a more efficient use of resources to enable desirable applications. AI with contextual awareness, AI explainability techniques or AI that can establish causality between decisions and their effects would all help to unlock several AI use cases that could bring revolutionary benefits to decision-making.

2. **Develop verification and validation (V&V) procedures for AI wargaming tools –**
 Experts highlighted the prevalence of non-AI automation being marketed as ‘AI’, and risks around inexperienced customers procuring low-quality AI-enabled wargaming tools. Standardised V&V procedures will be important to ensure the reliability of AI tools for wargaming and their compatibility with rules of engagement, ethical standards and explainability requirements.

3. **Trial narrow AI use cases in parallel to investing in complex AI-enabled wargaming –** There is a high cost to benefit ratio in the more sensationalised AI use cases such as CoA analysis and adjudication explored in this report. There is merit in testing implementation of narrow and relatively low risk automation and AI use cases (e.g. Procedural Content Generation, first pass scenario generation, and speech transcription) in wargames, , while simultaneously investing in research to advance complex AI-enabled wargaming (e.g. AI for player decision-making analysis and AI in adjudication) in parallel.

4. **Commission research on wargaming epistemology and decision-making in wargaming –** A better understanding of an effective wargame and human decision-making during a wargame would enable the design of AI-enabled tools that augment rather than add uncertainty to the validity of wargame outputs. In strategic wargames, a better understanding of quantifiable parameters that affect strategic outcomes would be required. The values that decision-makers apply in their decision-making would also need to be identified.

Appendix A. Long-list of AI-enabled Techniques

A longlist of AI-enabled techniques identified throughout this research project is provided below (as illustrated in Figure 5). For each use case, we describe the envisaged benefits and challenges, drawing on insights across the research interviews, workshop, and literature review. Finally, we give an indication of whether workshop participants saw each application as meriting prioritisation. Views on this topic were highly variable and so in many cases divergent views, alongside reasons for divergence, are given. These tables will provide a valuable reference point for anybody looking to learn more about a planned AI application for wargaming.

Figure 5. Long-list of AI use cases for wargaming identified.

Game design	Game execution	Game analysis	Game logistics
AI red team agents	Simulating opponents' reactions	AI to highlight interesting interactions between teams	Real-time language translation
Procedural Content Generation	Adjudication	Data capture on player behaviour	Collating past wargame data with OSINT and secret information
Scenario and inject generation	AI to run through tactical decisions	Player decision-making analysis	Speech transcription
Gameplay testing	AI decision assistants and CoA analysis tools		
	Virtual assistant (to provide legal reminders)		

A.1 Game design

Table 5. Summary of AI applications for game design.

Description of AI application	Envisaged benefits	Challenges	Rationale for (de)prioritisation ¹⁰⁸
AI red team agents (e.g. via agent-based modelling) ¹⁰⁹	Reduces number of personnel required; Varied red team characteristics can be tested; Games can be easily repeated	Reduces ability to study human decision-making (becomes modelling and simulation); ¹¹⁰ Training data frequently unavailable for realistic red-teaming; ¹¹¹ Emergent behaviours of AI agents are difficult to validate for realism; ¹¹² Actions of AI red teams are not explainable. ¹¹³	<p><i>Divergent views:</i></p> <p>Invest: Some favoured investing in AI red teams for more tactical games.</p> <p>Track: Many participants thought AI had potential to be integrated into red team activities in the future but cautioned that current models are not sophisticated enough to capture complex human behaviour. Since there are several commercial entities already working on this aspect, it will be better to track industry progress.</p> <p>Disregard: If the game itself is not digitised, it will be very challenging for AI to model the rules accurately. Additionally, modelling complex and unpredictable human interactions is not possible in the near term, especially for less quantifiable, strategic games.</p>

¹⁰⁸ Insights are drawn from a CETaS Workshop held on 16 March 2023. Where participants views were divergent, the order in which prioritisation rationales are presented herein is not indicative of which views were most popular.

¹⁰⁹ Author interview with academic wargame design expert, 30 January 2023; Author interview with government AI in wargaming expert, 26 January 2023; Author interview with government AI technical expert, 19 January 2023.

¹¹⁰ Author interview with government AI technical expert, 17 January 2023.

¹¹¹ Author interview with government AI technical expert, 17 January 2023; Author interview with private sector modelling and simulation expert, 15 February 2023.

¹¹² Author interview with academic wargame design expert, 18 January 2023; Author interview with government AI technical expert, 17 January 2023.

¹¹³ Author interview with academic wargame design expert, 30 January 2023.

Procedural Content Generation (e.g., via deep learning) ¹¹⁴	Increases speed of development of game mechanics; Improves game realism by adding in more detail to background content/non-player characters	Lack of control of generated content; ¹¹⁵ High cost-to-reward ratio outside of commercial sector with big budgets	<p><i>Divergent views:</i></p> <p>Deploy: AI can be already do this within limited scope (e.g. producing certain wargaming rules, developing non-player characters).</p> <p>Invest: While the simplest forms of PCG could already be deployed, more investment is needed to harness its full potential.</p> <p>Track: Some thought it possible to piggyback off of developments in industry.</p>
Scenario and inject generation (e.g., via generative AI) ¹¹⁶	Increases speed of generation of text and visual aids (e.g., DALL-E type tools) to the game to reduce time and costs; ¹¹⁷ Improves game realism and player immersion	Current techniques are only sophisticated enough for 'background content,' or as a first pass which requires further edits by a human; ¹¹⁸ third-party models are not always seen as sufficiently	<p><i>Divergent views:</i></p> <p>Deploy – Several participants cited already using tools like ChatGPT to produce a first-pass version of scenarios, cutting significant time and costs to scenario development. This can free up time to focus on more critical game elements.</p> <p>Invest – Participants mentioned that this type of input should be explored further by Defence to ensure that the creative elements of any content generation are robust. And, while a lot of content can be generated via simple automation already, investment in future advancements of this through AI were (broadly) seen as worthwhile to pursue.</p>

¹¹⁴ Jialin Liu et al., "Deep Learning for Procedural Content Generation," *arXiv* (October 2020).

¹¹⁵ Jialin Liu et al., "Deep Learning for Procedural Content Generation," *arXiv* (October 2020).

¹¹⁶ Author interview with think tank wargame design expert, 24 January 2023.

¹¹⁷ Author interview with private sector modelling and simulation expert, 1 February 2023.

¹¹⁸ Author interview with academic wargame design expert, 17 January 2023; Author interview with think tank modelling and simulation expert, 2 February 2023; Author interview with private sector modelling and simulation expert, 1 February 2023.

trustworthy;¹¹⁹ loss of creative richness in in content.¹²⁰

Game play testing (e.g., via agent-based modelling, artificial neural networks and/or deep learning) ¹²¹	Ensures a more comprehensive picture of possible outcomes of games; reduces risk of unexpected outcomes; helps identify unexpected mistakes in AI-generated content ¹²²	Expensive and time consuming for wargames taking place outside of simulated environments ¹²³	<p><i>Divergent views:</i></p> <p>Deploy: Provided that human designers have the ultimate say in finalising different game play elements simulated by AI, several participants believed that models could help to test a wider number of potentialities in games compared to humans.</p> <p>Track: Industry is ahead in this and there is a high risk that AI will not be able to model the game sufficiently to achieve this objective.</p> <p>Disregard: Other participants would disregard this use case for manual wargaming due to the high costs associated with creating a digital twin of the war game to facilitate testing using AI agents.</p>
---	--	---	--

¹¹⁹ Author interview (2) with AI technical expert, 11 January 2023.

¹²⁰ CETaS Workshop, 16 March 2023.

¹²¹ Author interview with government AI technical expert, 19 January 2023.

¹²² Daniel Karavolos, Antonios Liapis and Georgios Yannakakis, "Learning the patterns of balance in a multi-player shooter game," *Proceedings of the 12th International Conference on the Foundations of Digital Games* (August 2017): 1–10.

¹²³ Author interview (2) with AI technical expert, 11 January 2023.

A.2 Game execution

Table 6. Summary of AI applications for game execution.

Description of AI application	Envisaged benefits	Challenges	Rationale for (de)prioritisation ¹²⁴
Simulating opponents' responses (e.g., via predictive analytics) ¹²⁵	Reduces number of personnel required; novel opponent actions may be discovered	Lack of data on real-life behaviour patterns of opponents; difficulties in verification and validation of outcomes; explainability challenges.	<p><i>Divergent views:</i></p> <p>Deploy: Some thought that AI could identify a wider range of possible opponent reactions, thereby identifying new directions in which the game could go.</p> <p>Invest: Further research could help elucidate whether it is possible for AI to draw on historical outcomes to give probabilities on how opponents would react to specific situations. Warrants more exploration at tactical levels of warfighting and in deductive exercises because we have patterns of data, but less useful in strategic games.</p> <p>Track: Several participants felt more research was needed on how accurate AI predictions were on this aspect of the game, since large margins of error here would risk skewing other game dynamics (e.g., blue team responses).</p>

¹²⁴ Insights are drawn from a CETaS Workshop held on 16 March 2023. Where participants views were divergent, the order in which prioritisation rationales are presented herein is not indicative of which views were most popular.

¹²⁵ Author interview with government wargaming experts, 10 December 2022.

Adjudication (e.g., via logistic regression) ¹²⁶	Makes adjudication less random; allows integration of operational data or prior wargame data into future adjudication decisions	Not sufficiently flexible to account for subjective nature of adjudication; ¹²⁷ lack of data; ¹²⁸ not sufficiently explainable ¹²⁹	<p><i>Divergent views:</i></p> <p>Track: Scepticism about AI taking on this role is widespread. Industry projects should be tracked to understand if this is something that could be considered later down the line, specifically in the context of tactical games given limited potential at the strategic level.</p> <p>Disregard: Sometimes, adjudication is based on rules and mechanics. AI does not have a place here as simpler automation is possible. Otherwise, adjudication requires complex human judgement and understanding of the strategic context, as well as the possible military tactics and strategies that may be deployed. These decisions are best taken by humans with expert knowledge in these fields.</p>
AI to run through repetitive tactical decisions (e.g., via AI modelling and descriptive analytics) ¹³⁰	Increases player time available for more relevant/complex decisions. ¹³¹	Cumulative impacts of small errors can build up to skew overall game outcomes.	<p><i>Divergent views:</i></p> <p>Deploy: Some suggested this could already be deployed in narrow use cases.</p> <p>Track: Since many participants felt that AI's role here would not have too much impact on important decisions of the game (e.g., at the strategic-level), this could be a great way of freeing up players from mundane considerations.</p> <p>Disregard: Others suggested that decisions made over tactical aspects could in fact have cumulative impacts. As such, humans should still be responsible.</p>

¹²⁶ Jennifer McArdle and Eric Hilmer, "Effectively Integrating Technology into Wargames," *Interservice/Industry Training, Simulation, and Education Conference* (2022).

¹²⁷ Author interview with think tank wargame design expert, 12 January 2023.

¹²⁸ Author interview with think tank wargame design expert, 23 January 2023; Author interview with private sector modelling and simulation expert, 1 February 2023.

¹²⁹ Author interview with think tank wargame design expert, 24 January 2023.

¹³⁰ Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

¹³¹ Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

AI decision assistants and Course of Action (CoA) analysis tools (e.g., via reinforcement learning) ¹³²	Increases speed of player decision-making; decreases stress of decision-making	If the entire system is not digitised, it becomes difficult for AI to understand player moves and decisions.	<p><i>Divergent views:</i></p> <p>Invest: Defence research organisations are already investing here, and commercial industry will not. Defence has to develop this capability in-house.</p> <p>Disregard: Technically infeasible because it is difficult for AI to make sense of an environment with no rules.</p>
Virtual assistants that provide reminders on legal and regulatory framework, rules of engagement, updates on real-world policy events and	Reduces cognitive load on players; Improves ethical decision-making of players; improves player immersion. ¹³⁴	Slippery slope towards players playing a less active role in the game; ¹³⁵ lack of explainability contributes to distrust; ¹³⁶ likely to be a highly costly and technically complex development; ¹³⁷ establishing whose values, standards	<p><i>Divergent views:</i></p> <p>Invest: Given the decisions virtual assistants would be involved with, it is necessary to ensure they are tailored to the specific needs of defence stakeholders. Investment from within this sector is therefore needed. Participants see developing this capability as an enormous task, especially given the need to develop UK-specific tools for the UK legal context, but one with potentially significant benefits. It would be critical for any AI model in this capacity to start off with a more passive role supporting players (e.g., pointing out why they may have failed in winning a scenario) before gradually building up complexity. Doing so may also offer a useful training benefit where players can identify where they went wrong.</p>

¹³² Alec Barker, "BrainSTORM: exploring artificially intelligent COA development in STORM," *SPIE.Library* (April 2021).

¹³⁴ CETaS Workshop, 16 March 2023.

¹³⁵ Author interview with academic wargame design expert, 18 January 2023.

¹³⁶ CETaS Workshop, 16 March 2023.

¹³⁷ CETaS Workshop, 16 March 2023.

news analysis
(e.g., via
descriptive
analytics and
NLP)¹³³

should be incorporated
in ethical requirements
will be challenging.

Track: Currently, the level of complexity required for a role like this does not correlate to AI capabilities, meaning there is a serious risk that models could provide incorrect advice on important game mechanics (e.g., legal rules). Participants felt this is something which could benefit from monitoring industry developments in other fields (e.g., AI legal advisors), particularly since it is likely to be a highly costly and technically complex technique.

¹³³ CETaS Workshop, 16 March 2023.

A.3 Game analysis

Table 7. Summary of AI applications for game analysis.

Description of AI application	Envisaged benefits	Challenges	Rationale for (de)prioritisation ¹³⁸
AI to highlight interesting interactions between red and blue human agents to data scientists ¹³⁹	Reduces analysis time required; ¹⁴⁰ enables insights into tactical and strategic decision-making that would enable decision advantage	Limitations on AI's ability to identify what is salient mean important interactions may be missed; ¹⁴¹ AI could misinterpret true emotions or illogical actions. ¹⁴²	<p><i>Divergent views:</i></p> <p>Invest: A few participants advocating investing in this though further analysis was not provided.</p> <p>Track: Closely related fields such as diplomatic crisis management are making significant progress in this field and lessons can be learnt from them (e.g., use of AI to track the convergence of opinions during peace negotiations). However, this still needs more development and industry was seen as being likely to be exploring this anyway – so Defence should keep a close eye on developments.</p>
Data capture on player behaviour (e.g.,	Enables analysis of player decision-making	Wargaming becomes 'human subjects research' when such invasive data	<p><i>Divergent views:</i></p>

¹³⁸ Insights are drawn from a CETaS Workshop held on 16 March 2023. Where participants views were divergent, the order in which prioritisation rationales are presented herein is not indicative of which views were most popular.

¹³⁹ Author interview with government wargaming experts, 10 December 2022; Author interview (2) with think tank wargame design expert, 10 January 2023.

¹⁴⁰ Author interview with private sector modelling and simulation expert, 15 February 2023.

¹⁴¹ Author interview with think tank wargame design expert, 24 January 2023.

¹⁴² Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

via computer
vision)¹⁴³

collection is used,
bringing further
restrictions;¹⁴⁴
infrastructure is often
not there to facilitate
sophisticated data
collection;¹⁴⁵ difficult
to predict for a
particular game what
data are most
important to
capture;¹⁴⁶ invasive
data collection
protocols can impact
the way players
behave, introducing
unexpected game
outcomes.

Deploy: Since this is an output that is challenging for humans to replicate manually, utilising AI for this role could open up new insights into player decision-making that would help to enrich other analytical processes.

Track: Progress on computer vision is likely to be most advanced in other sectors and Defence can draw on this progress.

Player
decision-
making

Reduces possible
bias in players' self-
reporting of decision

May be less useful for
operational analysis
and more relevant to

Divergent views:

Deploy: While some participants placed this in 'deploy', they believed NLP can be used to support a human analyst, but the process certainly

¹⁴³ Author interview with think tank modelling and simulation expert, 17 January 2023.

¹⁴⁴ Author interview with think tank wargame design expert, 24 January 2023.

¹⁴⁵ Author interview (2) with think tank wargame design expert, 10 January 2023.

¹⁴⁶ Author interview with private sector modelling and simulation expert, 15 February 2023.

analysis (e.g., via NLP) ¹⁴⁷	logic; captures insights on why a move failed/succeeded quicker; generates insights on how players make decisions under stress; suggests avenues for reducing bias or novel strategic or tactical approaches; creates database on player decision-making to one day emulate a player. ¹⁴⁸	academic studies on wargaming; ¹⁴⁹ lack of transparency over how AI arrives at a judgment on player-decisions will likely create challenges in taking forward any AI-generated analysis. ¹⁵⁰	<p>can't yet be fully automated. Thorough testing would also be required to check the accuracy of these systems in the defence context.</p> <p>Invest: Participants saw utility in applying AI for emotion detection and analysis of player's behaviour to gather data and optimise decision-making performance under stressful conditions.</p> <p>Track: Many participants felt that there is not enough data or trials using AI solely for analysing human decisions in gaming, meaning it would need more R&D before considering integration. Additionally, there was scepticism over how game analysts would be able to understand the logic chain or rationale behind why AI arrived at a judgement on decisions, which not only risks reducing trust, but reduces the richness of findings.</p>
---	--	--	---

¹⁴⁷ Author interview with private sector modelling and simulation expert, 24 January 2023; Author interview with government AI technical expert, 19 January 2023.

¹⁴⁸ Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

¹⁴⁹ Author interview with think tank wargame design expert, 23 January 2023.

¹⁵⁰ CETaS Workshop, 16 March 2023.

A.4 Game logistics

Table 8. Summary of AI applications for game logistics.

Description of AI application	Envisaged benefits	Challenges	Rationale for (de)prioritisation ¹⁵¹
Real-time language translation (e.g., via NLP) ¹⁵²	Enables international collaboration and understanding of non-English discourse during the game	Challenges over AI capturing nuance, idioms, sarcasm or accents which could impact game decisions. ¹⁵³	Deploy: Seen as the easiest application to set up with significant value. However, there would be a need for thorough testing – in particular with regard to whether these technologies can handle 1) highly specialised military terms and their different usage across countries, 2) varied accents and 3) nuanced language use without introducing bias.
Collating wargame data with secret and OSINT, automating risk analysis (e.g., via	Facilitates increased collaboration across departments; increases speed of decision-making; enhances realism of game and re-introduce complexity of real-world decision-making in	Infrastructural and classification obstacles to this sort of data sharing;	<p><i>Divergent views:</i></p> <p>Deploy: Participants consider this collation to be technically possible now without any AI applications.</p> <p>Track: Participants were concerned over how trustworthy existing AI applications are in securing this confidential data against cyberattacks,</p>

¹⁵¹ Insights are drawn from a CETaS Workshop held on 16 March 2023. Where participants views were divergent, the order in which prioritisation rationales are presented herein is not indicative of which views were most popular.

¹⁵² Author interview (2) with think tank wargame design expert, 10 January 2023; Author interview with AI technical expert, 11 January 2023.

¹⁵³ CETaS Workshop, 16 March 2023.

artificial neural networks) ¹⁵⁴	game environment; strengthens evidence-base for critical decisions; could be applied to select customer requirements or questions to guide future wargames quicker. ¹⁵⁵		intrusions, etc. Nevertheless, they felt it would be useful to monitor any industry developments on this aspect.
Speech transcription (e.g., via machine learning) ¹⁵⁶	Improves completeness of data capture; ensures transcripts are available to store and use in future game design/development/analysis; reduces potential note-taker bias	Large quantity of data generated will often be of limited use and too time consuming to sift through; concern over how AI could capture nuance and accents in different languages. ¹⁵⁷	Deploy: Simplest form of transcription technologies are seen as ready for deployment with thorough testing. Other adjacent technologies, such as summary technologies, would need further work to ensure they are explainable. Risk of deployment is lower given the less technology interferes less with the players.

¹⁵⁴ Author interview with government AI technical expert, 17 January 2023.

¹⁵⁵ Author interview (2) with private sector modelling and simulation expert, 1 February 2023.

¹⁵⁶ Author interview with think tank wargame design expert, 24 January 2023; Author interview (2) with think tank wargame design expert, 10 January 2023; Author interview with former government wargame design expert, 18 January 2023; Author interview with think tank wargame design expert, 11 January 2023.

¹⁵⁷ CETaS Workshop, 16 March 2023.

Appendix B. Glossary

B.1 Wargaming lexicon

Wargame: For this study, we adopt the definition of wargaming used in the MoD Wargaming Handbook. Here, wargaming is defined as ‘a scenario-based wargame model in which the outcome and sequence of events affect, and are affected by, the decisions made by the players.’¹⁵⁸ This player-centred definition allows us to explore the ways AI might alter the role of the players.

However, wargaming is not a monolith. Many techniques and processes fall under this umbrella term or are closely related in some capacity. Further definitions are therefore provided below.

The wargaming landscape:

Table 9. Table displaying key wargaming definitions.

Term	Definition
Analytic wargame	Wargames ‘used to explore national-strategic, strategic, operational and tactical issues across the full spectrum of military activity.’ ¹⁵⁹ Analytic wargames seek to simulate and understand human decision-making processes of adversaries, allies and other key stakeholders, ¹⁶⁰ as well as seek to illuminate a central dilemma that commanders should make hard decisions about. For the purposes of this study, analytic wargames can be contrasted with training wargames and fully simulated models which are out of scope.

¹⁵⁸ HM Government, *Wargaming Handbook* (Ministry of Defence: 2017), <https://www.gov.uk/government/publications/defence-wargaming-handbook>.

¹⁵⁹ HM Government, *Wargaming Handbook* (Ministry of Defence: 2017), <https://www.gov.uk/government/publications/defence-wargaming-handbook>.

¹⁶⁰ Author interview with private sector modelling and simulation expert, 15 February 2023; Author interview with think tank wargame design expert, 10 January 2023

Manual wargame	Increasingly, wargaming occurs in simulated or partially simulated environments. For this study, we are most interested in wargames which occur outside simulated environments.
----------------	---

Modelling and simulation	A related analytic methodology where a model is used to represent a system or process and a simulation is used to execute this model over time to determine what the outcomes will be. This technique does not require human players.
--------------------------	---

B.2 AI lexicon

Artificial intelligence: For the purposes of this workshop AI will simply be defined as the automation of cognitive tasks which would typically have been the sole domain of human beings. Rather than focusing on AI in general, this study will wherever possible reference specific AI techniques which may be of use in segments of the wargaming workflow.

AI Techniques:

Table 10. Table displaying key AI definitions.

Term	Definition
Machine learning	Machine learning finds patterns in data and make predictions. Data are used to train a model from the input data and the model is updated as new data are fed in, allowing it to detect new patterns.
Supervised and unsupervised learning	During unsupervised learning, an algorithm is used to identify patterns in unlabelled datasets through techniques such as clustering and association. During supervised learning, labelled datasets, where the desired output from the algorithm is known, are used to train an algorithm.
Reinforcement learning	Reinforcement learning is another machine learning technique, where an algorithm's behaviour is learned through interactions with the environment. The developers must create a method for either rewarding desired outcomes or punishing negative outcomes to facilitate this learning behaviour.

Hierarchical reinforcement learning	Hierarchical reinforcement learning is a method used to address reinforcement learning challenges by breaking the problem down into a hierarchy of subtasks where the tasks at the base have lower computational complexity but can nevertheless feed into the top-level goal. ¹⁶¹
Neural network	Neural networks are machine learning models whose structure is derived from the human brain, consisting of an input layer, at least one hidden layer, and an output layer. They consist of multiple nodes, each of which will have an input and output that contributes to the overall functioning of the network. ¹⁶²
Agent-based modelling	An agent-based model is a computational model which attempts to represent human agents and their interactions with one another and their environment.
Descriptive, predictive, and prescriptive analytics	<p>Data can be analysed in multiple ways to gain insights. Three key categories for such analysis are:</p> <ul style="list-style-type: none"> • Descriptive: focused purely on data from the past and used to account for present circumstances. • Predictive: focused on using past data to predict future patterns. • Prescriptive: focused on using outcomes from descriptive and predictive analytics to predict what actions would be most favourable in a given context.
Fuzzy modelling	Fuzzy models offer a mathematical method which can be used to represent imprecise or vague information. This is possible because, while traditional programming deals in 1s and 0s through a binary logic, fuzzy logic allows for values in between.

¹⁶¹ Bernhard Hengst, "Hierarchical Reinforcement Learning," *Encyclopedia of Machine Learning* (2011): 495-502.

¹⁶² "What is a neural network?," IBM, no date, <https://www.ibm.com/topics/neural-networks>.

About the Authors

Anna Knack is a Senior Research Associate in the Turing Defence and Security programme and Lead Researcher at CETaS. Anna's recent and ongoing research is focused on human-machine teaming, AI explainability, AI-augmented decision-making and cyber AI. Prior to joining the Alan Turing Institute, Anna was Deputy Co-ord Lead of the Technology, Disruption & Uncertainty research workstream in RAND, where she coordinated a portfolio of work that leveraged futures methods to help policymakers understand and prepare for the future strategic operating environment.

Rosamund Powell is a Senior Research Assistant at CETaS. Prior to joining CETaS, Rosamund worked as an Ethics Research Assistant within The Alan Turing Institute, contributing to the public policy programme research agenda. While at the Turing, she has co-authored research reports exploring topics at the intersection of artificial intelligence, ethics, and policy.



Centre for Emerging Technology and Security

RESEARCH REPORT