MASS

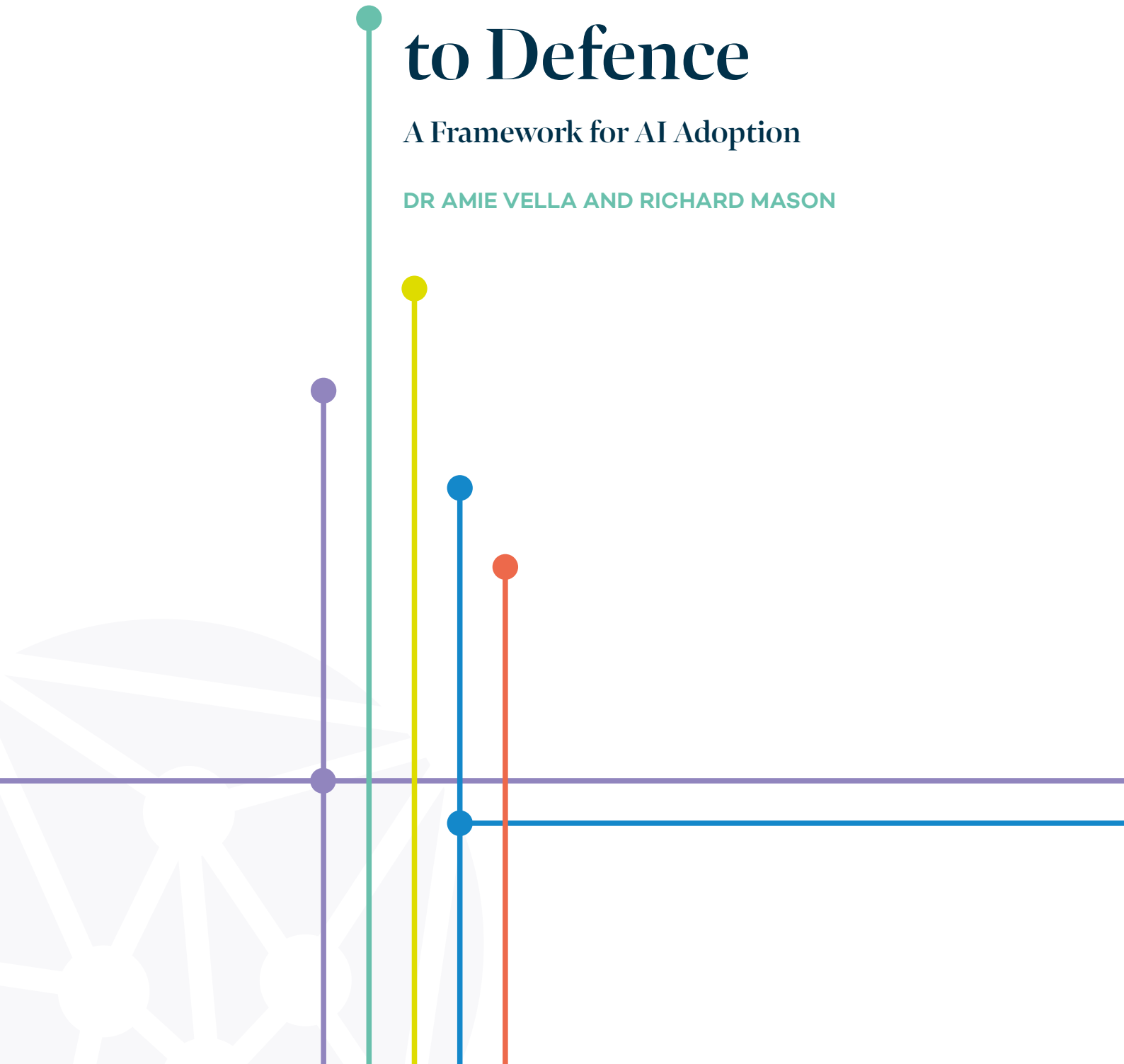
A COHORT PLC COMPANY

WHITEPAPER

Developing AI Solutions Tolerable to Defence

A Framework for AI Adoption

DR AMIE VELLA AND RICHARD MASON



MASS
A COHORT PLC COMPANY

About MASS

MASS is an independent, global technology company, trusted by highly-secure organisations to provide advanced, digital services that manage data and keep information safe. With our heritage in defence, we now offer robust solutions to sectors where security expertise is essential. We work in partnership with customers, using skilled, technical experts. We think innovatively to provide tailored, agile and resilient solutions that secure advantage, so you're ready for digital transformation.

PREPARED BY:

MASS Consultants Ltd
Enterprise House
Great North Rd
Little Paxton, St Neots
Cambridgeshire
PE19 6BN
United Kingdom

T: +44 (0)1480 222600

www.mass.co.uk

contact@mass.co.uk

VERSION HISTORY:

Version 1. October 2026

AUTHOR(S):

Dr Amie Vella and Richard Mason

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Contents

Executive summary 4

Abstract 5

Introduction 6

Generative AI considerations 8

Can AI technology help me to ...? 8

Figure 1 - The Hallucination Rate of some of the leading LLM's 8

Figure 2 - The Hallucination Rate of some of the leading LLM's 9

Hallucinations are not limited to AI 10

Figure 3 - Understanding the differences between challenging an AI hallucination versus a human one. 10

Factors for design considerations 11

The technology toolbox 13

Figure 4 – AI feature selection is dependent on the use case and data requirements 15

The social aspect conundrum 16

Competency development of AI assistants 17

Figure 5 – The competency development of AI versus a human equivalent operator 18

Policy, regulation and constraints 18

A Framework for AI Adoption in Defence 21

Figure 6 – Design thinking process that focusses on delivering outcomes in a structured collaborative way. 21

Discover: understanding decisions, context and risk 22

Figure 7 – The process starts with the Discovery phase 22

Define: establishing shared understanding, boundaries, and evidence 23

Figure 8 – The process moves to the Define phase, revisiting discovery to build understanding. 23

Develop: designing, challenging and learning before commitment 25

Figure 9 – Advancing to the Develop phase with a common understanding of what is required. 25

Deliver: building trust, integrating context and sustaining value 26

Figure 10 – The Deliver phase realises the outcomes but considers support & continual improvement. 26

Figure 11 – The DORA AI Capability Model 27

Cross functional framework success criteria 28

Figure 12 – How eliciting quality requirements spans Discover, Define and Develop phases. 29

Figure 13 – Aspects of test & evaluation across the framework aligned to the operational context of how it will be employed 30

Figure 14 – A cycle of continual monitoring, evaluation and system updates is necessary to maintain operational value. 31

From Socio Technical Risk to Structured Adoption 32

Learning from AI failures: why structured adoption matters 33

Table 1: AI Failures and how the Double Diamond approach could have helped 34

Conclusion 34

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Executive summary

AI offers Defence significant opportunities to improve efficiency, accelerate analysis and strengthen decision support. Yet adoption is not simply a technology challenge. Defence decisions are often high consequence, time-critical and shaped by complex operational, technical and human factors. Introducing AI successfully therefore requires a structured approach that balances innovation with assurance, accountability and trust.

This paper argues that Defence organisations should move beyond technology-led adoption. They should begin by understanding the problem, the decision being supported and the level of risk that can be tolerated. AI technologies have advanced rapidly, but they remain non-deterministic and can present inaccurate or misleading outputs with unwarranted confidence. The operational impact of these behaviours depends on the context in which AI is used and on the role people play in validating and acting on its outputs.

To address these challenges, MASS applies an AI Adoption Framework grounded in Design Thinking and Systems Engineering. It provides a structured route for organisations to explore opportunities, define requirements, experiment safely, generate assurance evidence and deploy AI responsibly. By integrating governance, testing, human factors and operational context throughout the lifecycle, organisations can realise value while maintaining trust, accountability and operational effectiveness.

AI can deliver substantial operational benefit to Defence—but only when applied to the right problems, with proportionate controls and clearly defined human oversight. The question is not whether Defence should adopt AI, but how it can do so safely, effectively and with confidence. MASS's application of the Double Diamond provides a practical, repeatable way to achieve that outcome.



© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Abstract

"AI-generated content may be incorrect." For Defence, that familiar warning raises a fundamental question: how much uncertainty is tolerable? Generative AI promises operational efficiencies, but those gains can quickly erode if people must spend significant time validating every output.

AI capability is advancing rapidly, and the case for adoption can appear compelling. The harder task is matching that technology to Defence processes, challenges and pain points, which differ markedly from those in other sectors. The Stanford University 2025 AI Index Report shows sharp improvements against human-performance benchmarks over the past two years, but without a structured implementation approach, organisations may fail to realise the benefits. During a recent hackathon with two NATO nations, we applied Generative AI to Electronic Warfare data analysis. The technology delivered valuable time savings, but it also produced mistakes and inconsistent results. Left unchecked, these errors could have serious consequences—reinforcing that not all decisions carry the same risk.

There are strong examples from which to learn. Google DeepMind's AlphaFold was built to solve a clearly defined problem: helping scientists understand how proteins interact with other molecules by predicting their complex three-dimensional structures. It has since predicted more than 200 million protein structures and helped researchers investigate proteins associated with disease. The central lesson is straightforward: DeepMind began with a specific outcome and designed the technology around it.

This paper offers NATO and Defence organisations a structured, collaborative framework for matching Generative AI capabilities to operational problems, drivers, outcomes and goals. It uses experimentation to build evidence, confidence and trust. The aim is to create solutions designed for specific Defence needs—reflecting the sector's distinctive accountability, criticality and operating constraints—rather than adapting technologies developed for other domains and hoping they fit.

Drawing on established Systems Engineering and Design Thinking methods, the framework demystifies the route to AI adoption. It helps organisations start with the capabilities that deliver value today while creating a disciplined path for experimentation, learning and continual improvement as the technology matures.

> **"AI's future isn't about bigger models. It's about smarter integration with people and processes."**
>
> **Jinesh Dalal**
> Head and Vice President, Technology Development, C-Metric

© MASS Consultants Ltd

Introduction

It is easy to be drawn to what AI appears able to do. Our experience shows that mature models can be valuable co-workers: summarising large documents, producing first drafts and supporting research into complex topics, with explanations and references where appropriate. Used well, these capabilities can save time and reduce routine workload.

However, our recent Electronic Warfare data-analysis hackathon with two NATO EW organisations also exposed important limitations. Generative AI outputs were not always accurate or consistent, reflecting the technology's non-deterministic nature and complicating testing and validation. Most notably, the models often responded confidently even when they were wrong, making errors harder to spot.

In Defence, such errors can have serious consequences. They also underline a central principle of this paper: not all decisions are equal. AI-generated content carries a different level of accountability and criticality when it informs operational judgement than when it supports low-risk administrative work.

Our experimentation revealed a practical trade-off between the ***time saved*** by Generative AI and the ***time spent*** checking its output. If validation demands outweigh the efficiency gained, adoption can increase operational burden rather than reduce it—and trust will quickly decline.

The Stanford University 2025 AI Index Report shows that performance against human benchmarks has improved sharply over the past two years and continues to advance. Defence should not avoid the technology; it should adopt it with controls that make each use case understandable, measurable and manageable. General benchmarks—such as multimodal understanding or reasoning scores—are useful indicators, but they do not capture Defence-specific operating conditions, decision types or risk tolerances. Bespoke experimentation and solution-level benchmarking are therefore essential to assess operational performance, reliability and suitability.



© MASS Consultants Ltd

One of the greatest challenges is that Generative AI often answers when it should acknowledge uncertainty. Independent measures such as the Artificial Analysis Omniscience Hallucination Rate show that even leading models can respond incorrectly rather than refuse or admit that they do not know. In Defence, a confident wrong answer is particularly dangerous because confidence affects whether an output is questioned, trusted or acted upon. This risk must be addressed through system design, context engineering, fine-tuning, tooling, evaluation, testing and integration.

Human long-term memory, learning, reasoning and inference remain indispensable. Current technology cannot match the depth of judgement built through experience. Artificial General Intelligence aims to deliver a more human-like level of capability, but it remains an unresolved scientific ambition. Human involvement is therefore not only an ethical requirement; it is operationally necessary because AI can reach the wrong conclusion.

For the short to medium term, we expect that the current technologies will continue to rapidly improve in their benchmark tests and this 'narrow form' of AI will continue to add value to the work we do. This is where we can realistically seek to address any specific use cases that present problems and use AI technology wisely. However, the outcomes achieved are shaped not only by model performance but by the social and organisational context into which the technology is introduced. The same AI system can produce very different results depending on how it is embedded within existing processes, hierarchies, time pressures, and cultures of trust.

This socio-technical reality demands a structured, problem-led approach to AI adoption. Defence must first understand the decision being supported, the tolerance for error and the role of the human in the loop; only then should it select or deploy technology. A clear definition of the problem, pain point or need for change is therefore crucial. Design Thinking provides a practical way to build that shared understanding and create a credible route to a solution.



© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Generative AI considerations

The number of Generative AI models available is increasing and their benchmark performance results are narrowing. There is a minefield of benchmarks to choose from, and it is not always clear how the results were derived, some using a Large Language Model (LLM) as a judge, with the risk of favouring their own answers & others using verifiable 'ground truth' answers such as Livebench.ai[1].

This section serves to surface some key factors when adopting Generative AI that should be considered in a system design to increase implementation success. When factored into the system design, these considerations will reduce the frustration of users and improve trust in the generated content.

Can AI technology help me to ...?

If you ask AI then, most often the response is yes. Ever noticed how when interacting with a AI agent by asking it if a particular task can be accomplished, it always answers ***positively*** and rarely says that it cannot help. This presents cause for concern, especially when the response answer is provided in such a convincing way. This highlights the need to ensure that where AI is being used in a business context, there needs to be a degree of AI literacy amongst employees and also a need to verify that the response is factually accurate.

By way of explanation into why AI responds so positively, benchmarking organisation Artificial Analysis[2] produces an AA-Omniscience Hallucination Rate table that shows how often a model answers incorrectly when it should have actually refused or admitted to not knowing the answer. Figure 1 below shows this benchmark result, where a **lower** value is better.

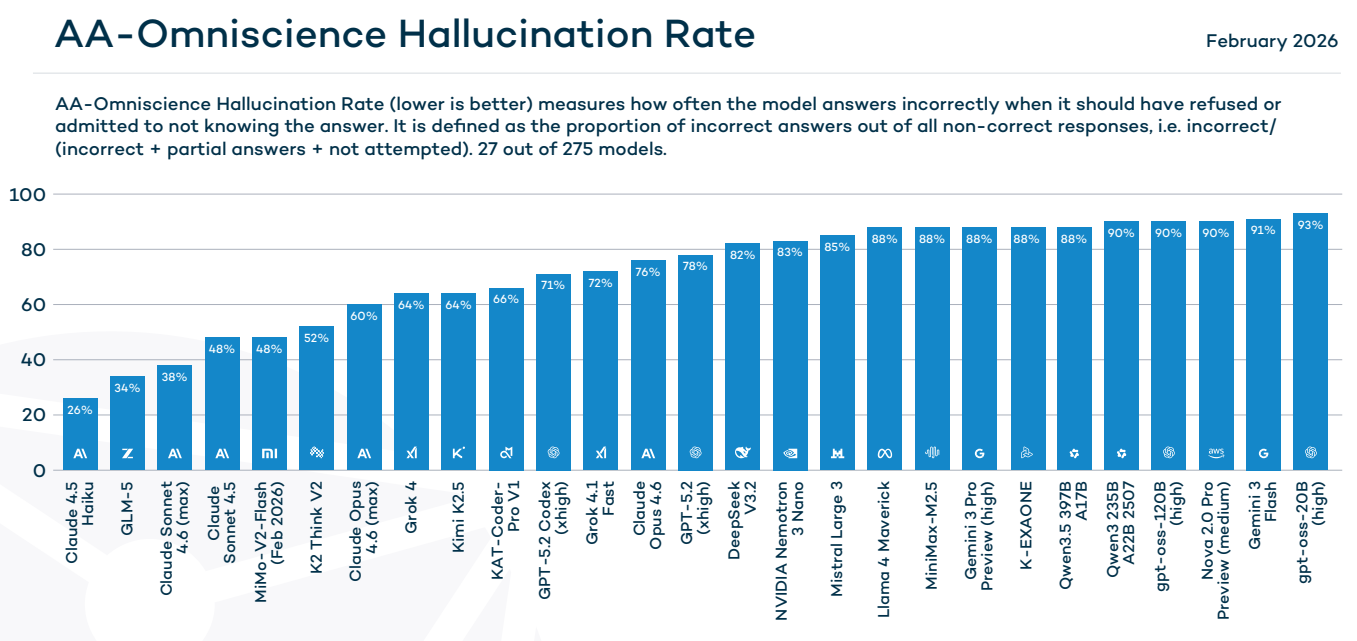


Figure 1 - The Hallucination Rate of some of the leading LLM's (Source: Artificial Analysis, February 2026)

1 (*LiveBench*, n.d.)
2 (*Comparison of AI Models across Intelligence, Performance, and Price*, n.d.)

As can be deduced from Figure 1, there is still a large proportion of potentially incorrect responses to a question asked to AI and this identifies the need for wise model selection in solution design, specific benchmarking on individual use cases and robust architecture development.

It further highlights that domain specific ***human knowledge*** is still an important part of employing systems that use Generative AI. Our ***long-term tacit knowledge*** based on our experiences over time is necessary to form an effective prompt, interpret AI generated results and help gain our confidence and trust in its use as a decision aid assistant.

Positive progress in improving hallucination shows in Figure 2, where in 6 months, the industry has shown to be addressing the issue with 13 models scoring below 40% versus just 3 as in Figure 1.

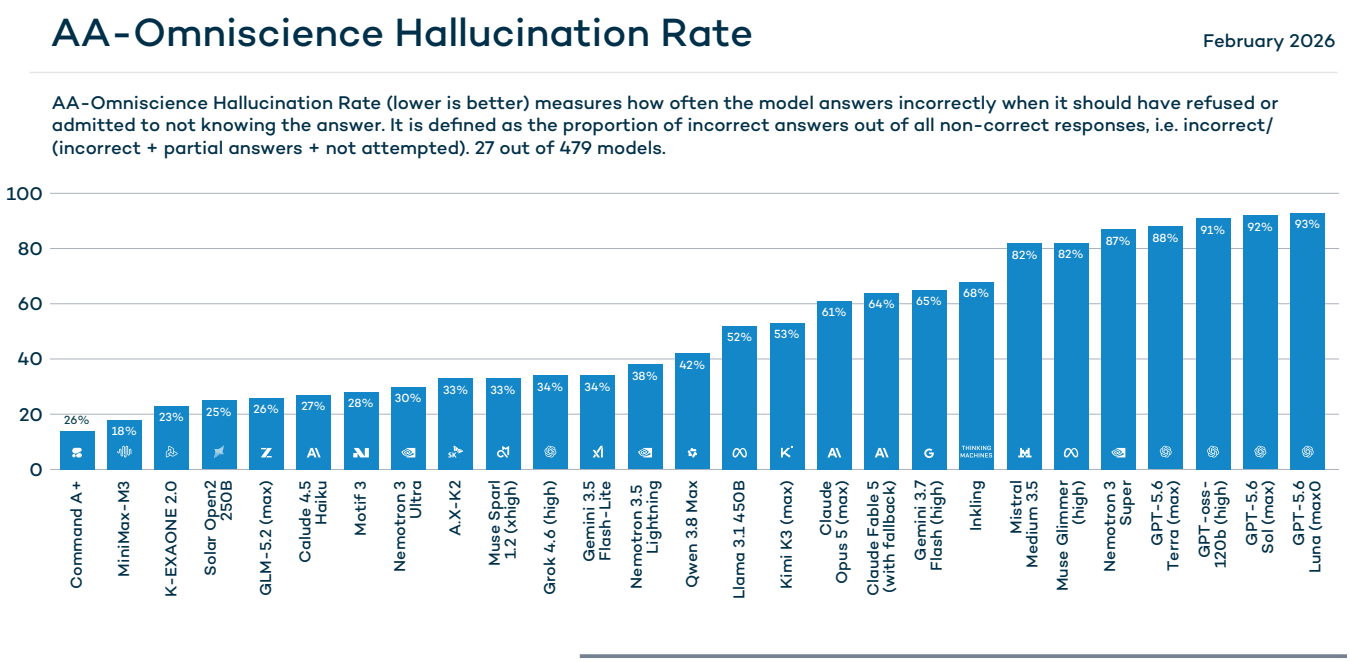


Figure 2 - The Hallucination Rate of some of the leading LLM's (Source: Artificial Analysis, August 2026)

Success consideration

Where a well-engineered system and a knowledgeable human operator, trained in its use converge, then there is scope for efficiencies to be realised in the use of AI.

MASS
A COHORT PLC COMPANY

Hallucinations are not limited to AI

An interesting observation to consider when faced with discussion around the hallucination aspects of AI is that ***actually***, humans hallucinate too. There is a subtle difference however when interacting with AI versus a human colleague, the human ***may*** get defensive when challenged, the AI system will not. Dependent on the personality traits of the human, they may be confidently sure of their assumptions and opinions, use their character to force opinion through, and it can sometimes be difficult to challenge their views with evidence; this can result in a prolonged time to reach the desired outcomes.

As a domain expert, using your judgement and tacit knowledge, you may notice that the AI response is incorrect or needs further elaboration. When you challenge the AI response, it will willingly work with you, through iteration, to achieve a desired outcome faster and this seems to be a consistent process, i.e. not impacted by human feeling.

Challenging AI is a useful tool in understanding the logic and thought processes it took, provides a good level of visible explainability and the resolution of a problem can be much faster than a human-to-human challenge.

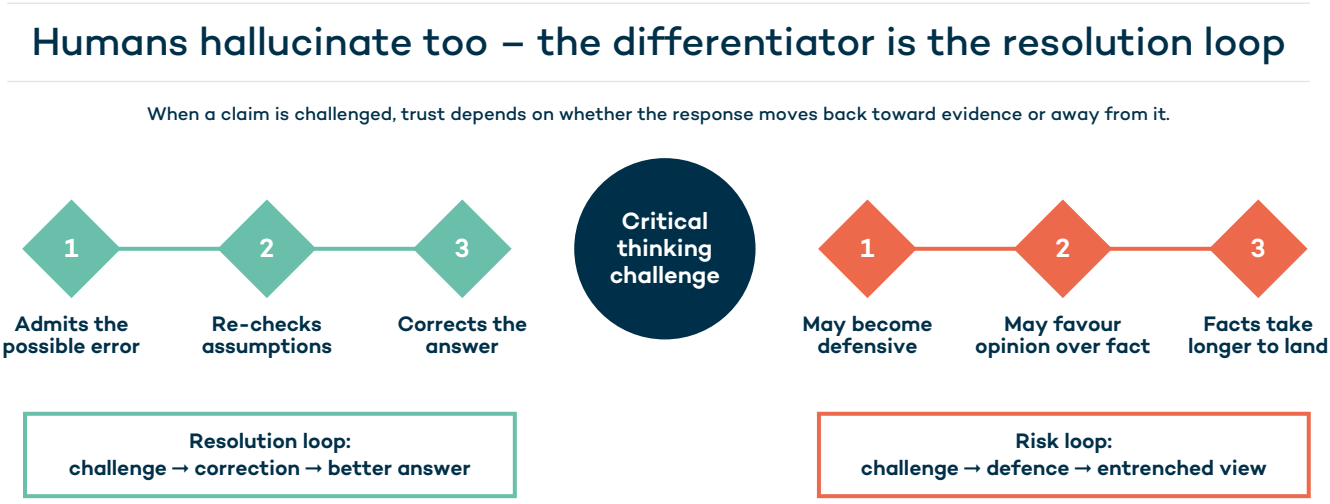


Figure 3 - Understanding the differences between challenging an AI hallucination versus a human one.

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Factors for design considerations

There are many variables that define a successful adoption of Generative AI, the quality of the prompt (Prompt Engineering), the correct choice of model and the quality of the data that the model is grounded in, i.e. the data & knowledge provided to the model (Context Engineering) to help improve the accuracy of results.

Defence terminology alone may require specific model training, as results of our own hackathon showed when performing Retrieval Augmented Generation (RAG) against the EW data within our THURBON® EW Data Management System. We learned that how people describe their request through natural language may differ, i.e. RADAR, Emitter, Sensor, Association, Fit, Fitted With, are domain specific language and therefore a key part in the specification, training and tuning of the system.

MODEL SELECTION

The growth in available models is both exciting and challenging. Whether that be a multi-modal model that can do a wide range of different tasks, a specific model focussed on a niche task or a range of models supporting a complex Agentic AI workflow. In fact, there is a lot of talk around LLM's where in fact, dependent on the use case, a Small Language Model[3] (SLM) may actually be a better fit, particularly for edge processing or workflows in Agentic AI.

Model selection for Defence applications may be constrained and dependent on the security classification, be hosted on secure cloud technology, a secure locally hosted model or on a capable edge device. These factors further highlight the need for elicitation of the problem we are trying to solve and ensuring that the right tools are selected for the individual use cases faced by Defence.

PROMPT EFFECTIVENESS

Effective prompts yield better results. Dependent on the experience of the person using the system, one may be happy with the results, another with less experience may be frustrated and lose faith and trust. To derisk the range of results, identification of common 'questions' that are needed can help to design effective 'pre-canned' prompts that produce more consistent and reliable results. Where the workflow is complex, multiple agents that ask the user questions can help steer and spread the workload to agents that have been designed and trained to do a specific task rather than a generalised prompt that the model struggles to comprehend in the correct context.

Success consideration

The quality of data and knowledge exposed to a Large Language Model is an important factor to be considered *and* discovering how a human uses this data in their work or decision process is helpful to ensure that access to the right data for the task is achieved.

3 (Belcak et al., 2025)

© MASS Consultants Ltd

CONTEXT AND ORGANISATIONAL KNOWLEDGE

Having the right quality data (knowledge) available to the system is crucially important to reduce hallucinations and increase the quality of the AI response. But it is actually more complicated than this. There are design choices to be made dependent upon the expected outcome of the use case. Each use case will be subtly different hence the need to really understand how we work and decide what technologies to employ in the system design that are the best fit.

For example, we may want to consider RAG to embed a large volume of key documents. This 'chunks up' the documents, embeds & vectorises them and stores these values in a database for the LLM to access via semantic search. Alternatively, we may want to directly attach a large document directly into the prompt to provide the context surrounding the question we are asking.

RAG was an engineering solution to the small 'context window' problem. The context window is the physical limitation on the size of the data that you provide to the AI prompt. It is the 'working memory' during the interaction and also determines the length of time that the LLM remains 'on topic' with the conversation. When using iterative prompting to achieve an answer, as we do in human conversation, the LLM can start to lose the thread of the topic and cause frustration to the user.

Similarly, RAG has limitations. Whilst the data is vectorised and stored, the information will be present, but there is a risk that it may not necessarily be retrieved when asked through a prompt. Naturally, advancements in the technology have greatly increased the context window size, and the prompt can be given much larger numbers of key documents to add context to the conversation session. Yet, this has a higher cost in the compute power needed.

EXPLAINABILITY

It will be no surprise that the explainability of AI is a concern across all domains and therefore a hot topic in the technology's development, and also in the field of Responsible AI. As humans we use a range of cognitive abilities and tools when reasoning and question our peers to reach understanding on how a decision was made. Whilst using AI, it is not unreasonable to expect the model to provide the rationale behind it's answer to a question, with citations to any reference material used and therefore give us greater confidence in the validity of the answer. In the simplest form, models support the 'explain your rationale' as part of a prompt, or as part of the system design to provide the 'chain of thought' process steps that the model executed. Understanding explainability and also accountability is a key consideration in the system design to provide evidence, where needed on the actions taken where AI was used as part of a decision-making process.

TESTING AND VALIDATING

We need to design systems, test and validate them such that they produce consistently reliable and expected results. This is a key consideration as part of the optimisation and tuning process to help reduce hallucinations, protect against bias or harmful content, and assess the quality, accuracy and relevance of non-deterministic outputs. It is key to understand how the system will be used in the operational workflow, alongside human team members and the decisions it will support. Developing a representative set

© MASS Consultants Ltd

of near real-world operational scenarios and decision vignettes early in the lifecycle can provide a realistic basis for testing and evaluation, allowing expected behaviours, performance measures and acceptable levels of risk to be established. These scenarios provide a repeatable mechanism for assessing system efficacy against meaningful operational outcomes rather than purely technical benchmarks and are explored further within the framework section.

Test & validation is a core part of the Software Development Lifecycle and should preferably be conducted by test teams trained in the International Software Testing Qualifications Board (ISTQB[4]) best practice. The test analysis, design, specification, test plan and tools used will help to de-risk the system delivery and provide the evidence that the system design is fit for purpose.

TECHNOLOGY MATURITY

As the technologies are at different levels of maturity, we can use the aspects of AI that work very well today, such as summarisation, image recognition, drafting presentations or executive summaries, translation or document comparison to produce efficiencies in our daily workflow. For it to be successful, particularly for complex use cases requiring clear accountability, we need a structured implementation framework that targets specific problems or pain in those workflows, considers all the variables and is designed to build trust in the system. In contrast, a system that doesn't follow this approach may in fact lead to increasing overheads in your workflow, reduced trust, increased cost, decreased reliability and a missed opportunity.

Success consideration

As can be seen, there is a complex web of factors that require consideration for specific use cases, highlighting the need for a structured implementatino framework.

The technology toolbox

At an organisational level, it is not necessary to fully understand what all the AI technology capabilities are, this is the role of the solution designer as part of the engineering process to map technology to specific use cases. In fact, the advancement and improvement of AI technology is changing on an almost weekly basis, and we can take comfort in this and look forward to working with the tools in our daily workflow. It is also not wrong to decide against the use of AI in a project. It may be that the technology is not yet quite the right fit and this can be revisited at a later date.

The time taken to do this up-front discovery and definition of the individual problems and pains encountered in the 'as-is' landscape will reap reward and allow solution designers to pinpoint exactly where AI technologies could be realistically integrated into the workflow. Additionally, this will greatly assist

4 (*International Software Testing Qualifications Board (ISTQB)*, n.d.)

in the definition of the experiments to test those theories, in conjunction with the domain experts experiencing the pain leading to a more viable and trusted resultant solution architecture. Equally, this process may identify areas where AI is not currently the most appropriate solution and alternative approaches should be considered.

Success consideration

Considering that we are focusing on applying the technology to the Problem *rather than asking* "how can we use the technology?", this puts great mphasis on the need for an integration framework to ensure that we capture how we currently carry out tasks enabling the matching of the right technology more effective downstream.

TECHNOLOGY SELECTION APPROPRIATE TO THE BUSINESS REQUIREMENT

Figure 4 serves to show that the technical solution to a problem or pain point is directly related to its specific operational use case. The selection may be just a single feature, or a combination of features and will require access to the data needed to fulfil the intended operational workflow activity.

The Context Layer provides the data, which will naturally differ between use cases, and in all instances Identity and Access controls are necessary to ensure that AI features are only explicitly accessing data appropriate to the task.

The AI Features Layer shows example technology features that can be selected during solution design and these may require the Orchestration Layer to co-ordinate the control of outputs across the solution such as connecting to external services (I.e. geospatial services), other data sources and applications.

The Governance Layer is applicable to all solutions, regardless of use case and ensures that any AI solution is given the regulation, guidance, evaluation, monitoring and support necessary to be part of an operational workflow; as we would do of course with an equivalent human team member.

© MASS Consultants Ltd

Mapping AI technology to specific use cases

The solution outcome will be a single, or combination of technology features with access to specific data dependent upon the operational use case.

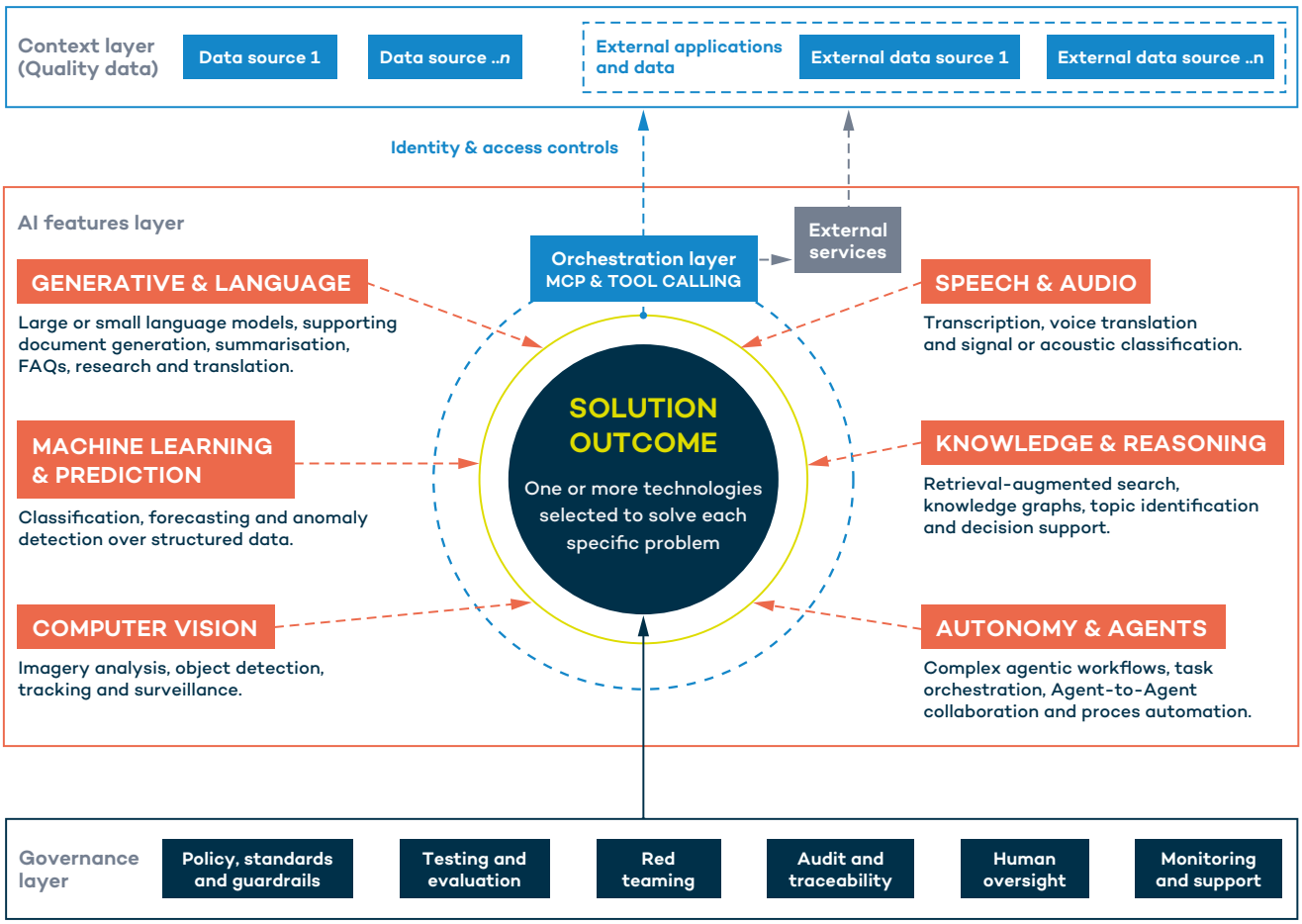


Figure 4 – AI feature selection is dependent on the use case and data requirements.

UNDERSTANDING WHERE THE HUMAN FITS

Determining where the human fits within an AI enabled workflow is not a secondary design choice but a core architectural decision. Given the tendency of AI systems to produce confident but occasionally incorrect outputs, human judgement remains essential – particularly where decisions carry significant operational consequence.

Human involvement should be explicitly designed based on the criticality of the decision being supported. For lower risk tasks, such as summarisation or exploratory analysis, humans may act primarily as consumers of AI output with light verification. For higher risk applications, humans may need to remain firmly "in the loop", responsible for validation, interpretation, and final decision making to help guide effective decision-making.

© MASS Consultants Ltd

This design choice must also account for socio technical factors. Under time pressure or within hierarchical decision structures, AI outputs may be afforded undue authority if their role is not clearly defined. Explicitly articulating the responsibilities of both the AI system and the human operator helps prevent automation bias and reinforces accountability.

Placing the human appropriately within the workflow also supports the retention of domain expertise and tacit knowledge. Rather than replacing judgement, AI should be designed to augment it – supporting sense making, highlighting patterns, or reducing cognitive load, while preserving the human's role in contextual reasoning and ethical responsibility.

The social aspect conundrum

The adoption of Generative AI in Defence cannot be understood solely through the lens of model performance, accuracy metrics, or computational capability. While these technical characteristics are important, they represent only part of the risk landscape. The outcomes achieved through AI-enabled systems are shaped equally by the social, organisational, and operational contexts into which the technology is introduced.

Generative AI systems do not operate in isolation; they are embedded within existing workflows, decision-making structures, cultures of trust, and organisational hierarchies. How outputs are interpreted, questioned, or acted upon is influenced by factors such as time pressure, perceived authority of the system, accountability structures, and the experience of the human operator. In Defence environments – where decisions are often time-critical and consequences significant – these contextual factors play a decisive role in determining whether AI-generated information supports or undermines operational effectiveness.

Recent academic research increasingly describes AI as a sociotechnical rather than purely technical problem. Kudina and van de Poel[1] argue that AI outcomes emerge not just from the technology itself but from the wider systems of people, processes, governance structures and organisational contexts in which it operates. Similarly, a large-scale review of AI adoption literature by Desveaud and Bawack[2] found that successful adoption is shaped by interactions between technology, users and socio-cultural context rather than technical capability alone. This reinforces the view that AI effectiveness cannot be separated from the environment in which it is deployed, governed and used.

This perspective aligns with socio technical thinking and the principles of the Social Construction of Technology (SCOT), which emphasise that technology does not determine outcomes on its[3]. Instead, outcomes emerge through the interaction between technical artefacts and the social systems that adopt and use them. From this viewpoint, risk does not reside solely within the AI model but emerges from how the technology is understood, trusted, and integrated into human decision-making processes.

1 (Kudina & van de Poel, 2024)
2 (Desveaud & Bawack, 2025)
3 (Bijker et al., 1987)

© MASS Consultants Ltd

A key challenge posed by Generative AI is its tendency to present responses with high confidence, even when the underlying information may be incomplete or incorrect. Confidence strongly influences behaviour and confident outputs are much more likely to be trusted, less likely to be challenged, and more likely to be acted upon – particularly under operational pressure. In Defence contexts, this creates a socio technical risk where errors are not only technical failures but behavioural ones, shaped by how humans interact with the system.

The same AI capability can therefore produce very different outcomes depending on where and how it is deployed. An AI tool used for exploratory analysis or administrative support may present minimal risk, whereas the same tool used to inform operational judgement may carry unacceptable consequences if its limitations are not well understood. This reinforces the principle that not all decisions are equal in their consequence, and that AI suitability must be assessed in relation to the decision being supported rather than the capability of the technology alone.

There is also a longer-term consideration regarding skill retention and tacit knowledge. As AI tools become more capable, there is a risk that over-reliance may erode the development of independent judgement and deep domain expertise. While AI can accelerate analysis and surface patterns, it does not replace the experiential knowledge that enables practitioners to recognise anomalies, contextualise information, and understand second-order effects. Maintaining this balance requires deliberate design choices that preserve meaningful human involvement rather than default automation.

For these reasons an effective AI adoption in Defence must treat socio-technical considerations as part of the problem definition, not as an afterthought. Understanding who will use the system, under what conditions, with what incentives, and with what tolerance for error is as important as selecting the right model or architecture. Trust, accountability, and human judgement must be designed into the system from the outset.

This socio-technical reality underpins the need for a structured, problem-led approach to AI adoption. Rather than starting with the technology and seeking applications, Defence must begin by understanding the decision context, operational environment, and human factors involved. Only then can AI be applied in a way that delivers efficiency while preserving trust, accountability, and operational effectiveness.

Competency development of AI assistants

Where AI tools are being used in a team, it can be useful to consider how the competency of the AI elements will be managed. With a human operator, we absolutely invest in their development through appointing levels of responsibility, formal and workplace training, provide Standard Operating Procedures, supervise and carry out performance reviews. We have the opportunity to treat and measure AI in much the same way, as shown at Figure 5.

© MASS Consultants Ltd



Figure 5 – The competency development of AI versus a human equivalent operator.

Policy, regulation and constraints

Any application of generative AI within defence and security contexts must be developed within a clearly understood policy, regulatory, and governance environment. These constraints are not external to system design; they actively shape where AI can be applied, how it must behave, and the degree of autonomy it can be granted.

Organisations should ensure alignment with existing defence AI strategies and ethical frameworks, which emphasise human centricity, accountability, and compliance with domestic and international legal obligations. In particular, human responsibility for decisions must remain clear and auditable, with AI positioned to support, not displace, human judgement in decision critical contexts. By applying the Design Thinking double diamond, these contexts are better considered.

Regulatory and policy constraints also extend to data governance, security, and access control. Classification boundaries, data provenance, and handling requirements place practical limits on model training, deployment architectures, and integration with operational systems. These constraints should be treated as design parameters rather than barriers, influencing architectural choices, evaluation strategies, and acceptable use cases from the outset.

© MASS Consultants Ltd

Finally, policy considerations should be revisited continuously rather than treated as a one time compliance activity. As AI systems evolve through learning, adaptation, or integration with new data sources, organisations should reassess risk, assurance, and alignment with policy intent. This ongoing governance is essential to maintaining trust, legitimacy, and operational confidence over time. The double diamond is well suited and relies on continuous improvement and an agile design process which suits the needs of a disruptive and ever-growing technological application, such as AI, as the needs of the technology and organisation adapts and learns.

This perspective is increasingly being reflected in international governance standards. The NIST AI Risk Management Framework[4] identifies trustworthiness characteristics including reliability, safety, explainability, resilience, accountability and bias management throughout the AI lifecycle. Equally, ISO/IEC 42001[5] positions AI governance as a management-system challenge requiring leadership, planning, operational control, continual evaluation and organisational learning. Together these frameworks suggest that successful AI adoption requires disciplined governance and risk management rather than focusing solely on technical implementation.

DEFENCE AI ASSURANCE AND RESPONSIBLE AI

Defence organisations are not starting from a blank sheet when considering the adoption of AI. Across MOD, NATO and allied nations, significant effort has already been invested in developing policies, principles and governance frameworks that seek to ensure AI is deployed safely, ethically and with appropriate human oversight. These frameworks recognise that the challenge is not simply whether an AI capability can perform a task, but whether it can be trusted to do so within complex operational environments where decisions may carry significant consequences.

The Ministry of Defence (MOD) and NATO have both adopted principles that have centred on human responsibility, accountability, reliability, traceability and governability. Whilst terminology may differ between frameworks, the common theme is the need to ensure that humans remain responsible for the outcomes generated by AI-enabled systems. This is particularly important in sectors like Defence, where operational effectiveness must be balanced against legal, ethical, safety and security considerations.

However, governance principles alone do not provide an implementation roadmap in which to follow. Organisations must still determine how to translate policy intent into practical system design, assurance evidence, testing strategies and operational procedures. This is often where adoption programmes encounter difficulty and there can be a tendency to focus assurance activities on the end of a project, when many of the key design decisions have already been made and are expensive to revisit.

This is why AI assurance should be viewed as a continuous activity that spans the entire capability lifecycle, from the beginning of a project and throughout. Assurance is not a tick box exercise; it how confidence is established and maintained. It requires organisations to understand the operational problem being solved, the decisions being supported, the potential consequences of error, and the controls necessary to manage risk at an acceptable level.

4 (Tabassi, 2023)
5 (*ISO – ISO 42001 Explained*, n.d.)

A structured adoption framework helps to create this assurance, and we will view this through the lens of the Double Diamond later on.

Importantly not all AI-enabled activities require the same level of assurance. The assurance associated with an AI assistant drafting administrative reports is significantly different from that associated with an AI capability supporting operational planning or intelligence analysis. Understanding decision criticality therefore becomes a key part of determining the level of human oversight, testing rigour, governance controls and evidence required before deployment.

The MASS AI Adoption Framework is not intended to replace existing Defence AI governance frameworks. Instead, it provides a practical mechanism through which the intent of Responsible AI principles can be translated into structured project activities. By embedding governance, assurance, experimentation, test and evaluation throughout the Discover, Define, Develop and Deliver phases, organisations can build confidence incrementally and ensure that AI capabilities remain aligned with operational needs, risk appetite and human accountability.

Successful Defence AI adoption requires governance frameworks and implementation frameworks to operate together. Governance establishes the principles and constraints within which AI should operate; implementation frameworks provide the processes, evidence and assurance activities necessary to demonstrate that those principles have been achieved in practice.

FROM COMPLEXITY TO CLARITY

Given the many factors discussed around successful AI adoption, including the technology considerations, behaviours, limitations and socio-technical factors it seems appropriate to seek assistance when considering use of AI in a defence organisation.

To enable this, we have produced a Framework for AI Adoption in Defence that helps organisations to pinpoint specific problems that could be improved with AI and a structured path to realising their vision and goals.

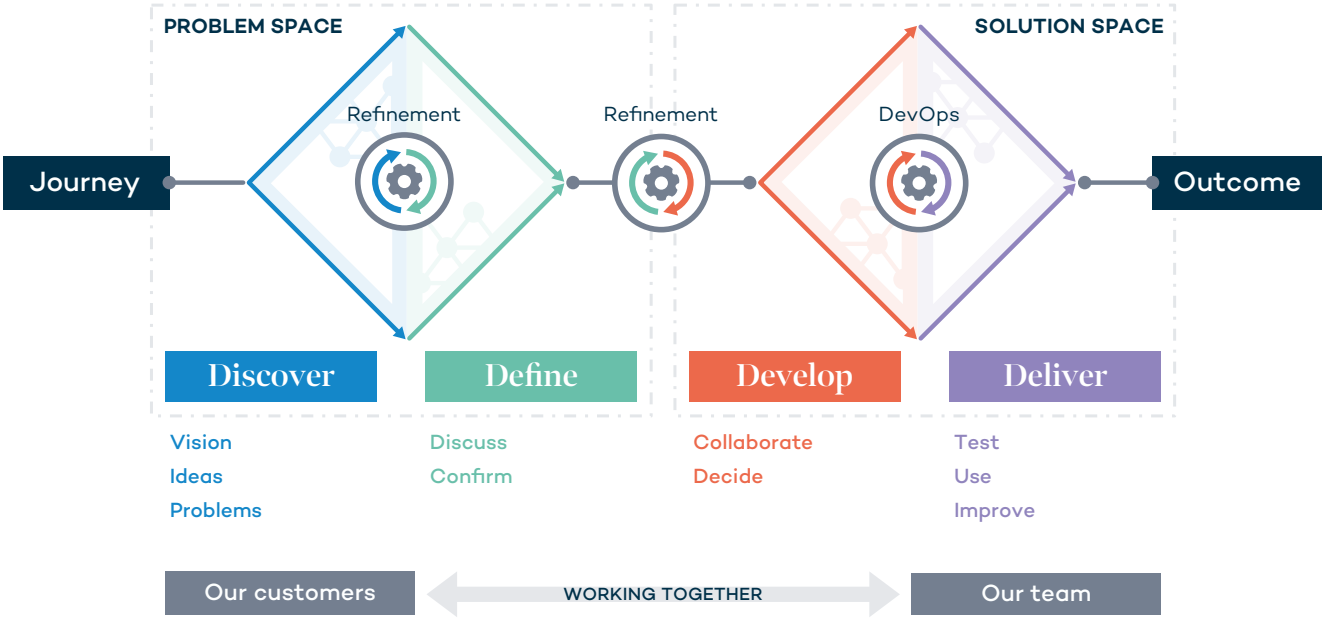
© MASS Consultants Ltd

A Framework for AI Adoption in Defence

The MASS AI Adoption in Defence Framework is outcome-focused. It creates a structured process for establishing the ground truth, clarifying problems, aligning vision and goals, and determining how best to achieve the required outcomes. At the outset, the solution—AI or otherwise—need not be discussed. The priority is to elicit robust requirements and give system architects the evidence they need to evaluate the right options later.

The process is based on the Design Thinking Double Diamond Framework and iterates through the 4 key phases of Discovery, Definition, Development and Delivery[6].

MASS' framework for AI adoption in defence



- Current situation
- Problems & pain
- Goals & objectives
- User journey
- User experience
- Contraints & risks

- Scope & playback
- Analyse & discuss
- Capture requirements
- Explor opportunities
- Experiment & learn
- Common understanding

- System engineering
- Investigate options
- Risk & cost driven architecture
- Collaboration
- System & UX design

- Agile build
- Test & validate
- Deliver & integrate
- Train
- Continual improvement
- Support

Figure 6 – Design thinking process that focusses on delivering outcomes in a structured collaborative way.

6 The Double Diamond - Design Council

© MASS Consultants Ltd

A central principle of the framework is that AI is not assumed to be the answer. The process first clarifies the problem, operational context, constraints and desired outcomes. Discovery, definition, experimentation and evaluation may show that deterministic software, conventional automation, process improvement or a change in ways of working is better suited to the need. AI can be revisited as the technology, governance and operational understanding mature. This is therefore not simply an AI implementation framework; it is a problem-solving framework that may conclude that AI is not the right solution.

The benefit of the framework is that it provides an agile structure to complex environments, offers the opportunity to circle back between each stage to refine aspects of both the ***Problem Space*** and the ***Solution Space*** and crucially provides both customers and suppliers with a common understanding of what is to be delivered.

As a project moves through each phase, the operational problem becomes clearer and the solution options more credible. The transitions between phases create deliberate opportunities to pause, test assumptions and reconsider whether AI remains appropriate. The organisation may conclude that the operational risk is too high, the technology is not yet reliable enough, or a deterministic solution would perform better. It may also decide to redesign the process and revisit AI when its maturity better matches the need.

At each stage, we use tried and tested tools and techniques to support the phase development including User Experience and Business Analysis through to Systems Engineering, Software Engineering and Architecture Design and System Testing. A description and purpose of each phase is outlined in the following sections.

Discover: understanding decisions, context and risk

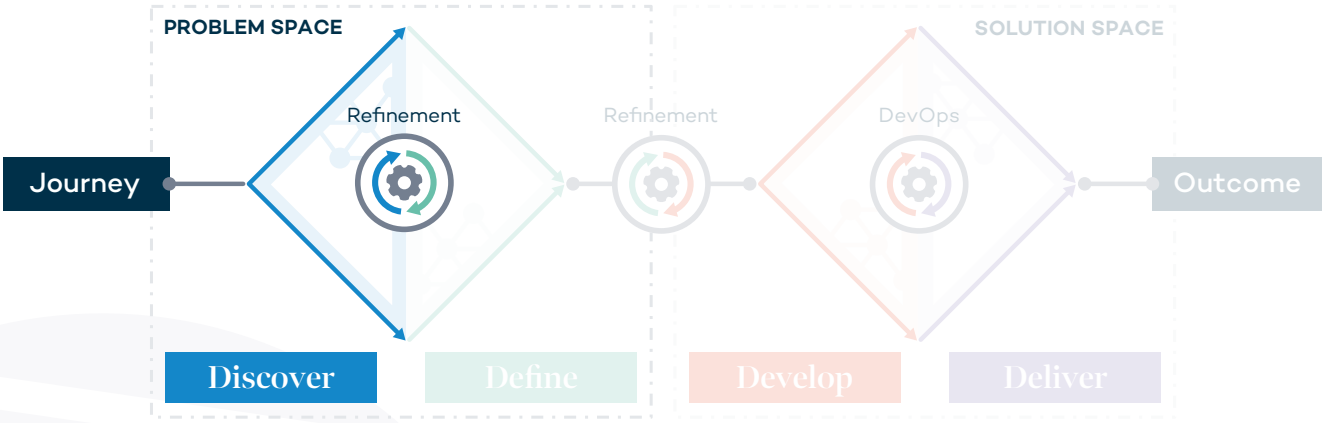


Figure 7 – The process starts with the Discovery phase.

When applying generative AI, organisations should begin with a discovery phase focused on understanding how decisions are currently made, rather than on the technology itself. This phase should prioritise listening and observation across the problem space in order to surface where uncertainty exists, how information is interpreted and where confidence or doubt shapes behaviour.

© MASS Consultants Ltd

Key activities at this stage include examining how decisions are taken in practice, not just how they are described in process documentation. This includes understanding what information people trust, what they routinely ignore or double check, and where information is incomplete, ambiguous, or delayed. These observations are critical for identifying where generative AI may genuinely support sense making, and where its involvement could introduce new risk.

Structured curiosity should then be applied to question and identify hidden complexity. Organisations should distinguish between what information *should* be used and what is *actually* used, identify signals that indicate uncertainty, and surface tacit or experiential knowledge that may not be formally captured. Particular attention should be paid to decisions that appear routine but carry disproportionate consequences if wrong or late, as these are often the most sensitive to automation.

In parallel, this phase should actively explore constraints and risks. Regulatory and policy limitations, organisational factors such as skills, trust, and accountability, and security boundaries including classification and access controls should all be made explicit. Understanding the consequences of incorrect or delayed decisions is essential at this stage, as it shapes whether AI can be safely introduced at all.

Discovery should converge through reporting and sense making, resulting in a shared problem narrative and an initial decision map. This should articulate areas of acceptable and unacceptable risk and form an early hypothesis about where AI might *support*, but not replace, human judgement.

Define: establishing shared understanding, boundaries, and evidence

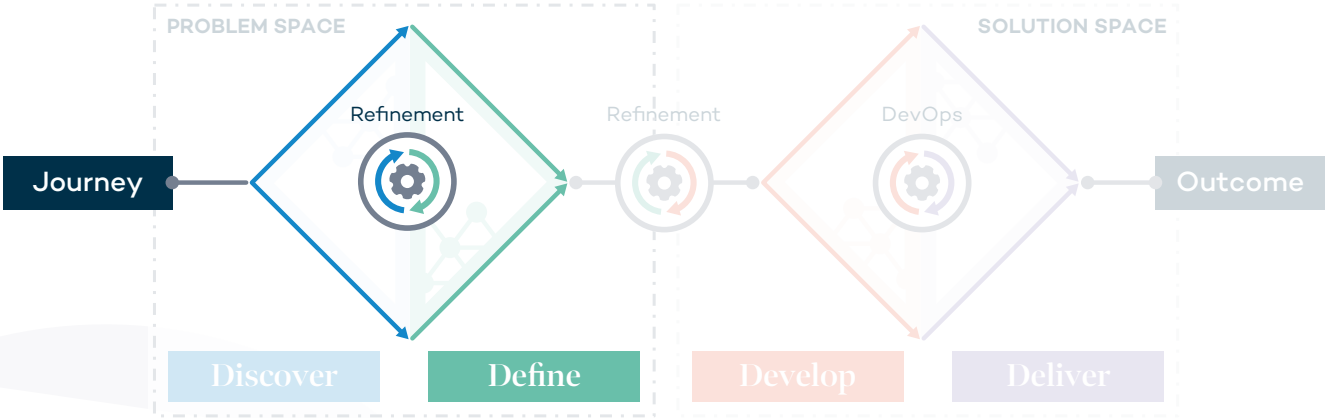


Figure 8 – The process moves to the Define phase, revisiting discovery to build understanding.

The define phase should be used to transform exploratory insight into a clear, shared, and bounded understanding of the problem that generative AI is intended to support. This phase is critical for preventing solution led adoption and for ensuring that AI use is intentional, constrained, and defensible.

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Success consideration

Success is not that a AI solution is specified in detail, but that the right probelm, boundaries an evaluation criteria are fixed early enough to prevent technical issues and operational failure.

Organisations should begin by discussing and confirming a shared problem narrative, agreeing which decisions are in scope and which are not. This includes defining what success looks like, clarifying decision ownership, and stating where AI is permitted to assist and where it should not be applied. Establishing these boundaries early reduces ambiguity and prevents later issues in system purpose or authority.

The problem should then be translated into requirements through targeted research and comparison. This includes examining what has worked elsewhere, understanding the constraints and failure modes of comparable solutions, and defining what the system must do, and how well, to be considered effective.

Requirements should encompass both functional and non-functional needs, including assurance, explainability, reliability, and trust.

At this stage, organisations should explicitly identify the knowledge sources that will provide context to the system and to its evaluation. This includes determining which data sources, reference materials, domain knowledge, and human expertise are authoritative, how these sources will be accessed or represented, and where gaps or bias may exist. Making knowledge sources explicit is essential for understanding how AI outputs are shaped and for avoiding misplaced confidence in de contextualised responses.

Success consideration

Success is about translating the discovery phase into clear and enforcable intent, it's about defining what success actually means, what boundaries exist and how risk will be adequately judged before a solution is developed.

The define phase should also establish evaluation strategies that replace opinion with evidence. Rather than assuming AI capability, organisations should define how success, failure, and uncertainty will be measured. This includes determining appropriate evaluation criteria, acceptable levels of error, and how performance will be assessed against real decision contexts. Evaluation strategies should be designed to surface model limitations and to support trust calibration, rather than to demonstrate best case performance.

The define phase should report and sense make and turn research and experimentation into refined decisions. Trade offs should be made explicit by agreeing requirements, validating the risk posture, and articulating early architectural implications. These decisions should be backed by evidence, providing a robust foundation for subsequent design and development activity.

© MASS Consultants Ltd

Develop: designing, challenging and learning before commitment

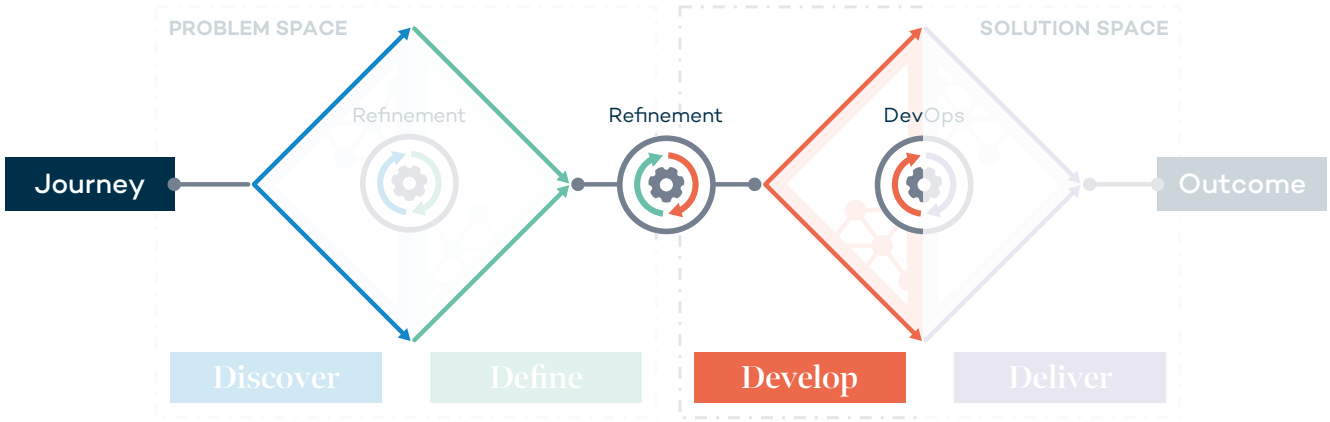


Figure 9 – Advancing to the Develop phase with a common understanding of what is required.

In the develop phase, attention should shift to how a AI solution would be designed, governed, and integrated, rather than whether it can function. Architectural decision making is central, including clarifying where AI sits within the system, such as decision support, drafting, or analytical assistance, and how it interacts with data and users as part of user accessibility and experience.

Quality attributes such as security, reliability, auditability, and risk management should be addressed explicitly. Development should include creating test plans, applying sound systems and software engineering principles, and designing solutions in a way that allows behaviour to be observed and interrogated over time, including for misuse or unintended consequences.

Intentional friction should be introduced through design reviews and challenge processes. These reviews should interrogate assumptions about trust and risk, clarify responsibilities and interfaces, and prevent the selection of technically interesting solutions that do not meet the operational need. Product breakdowns and design artefacts should make accountability visible rather than implicit.

Demonstration and collaboration are essential in this stage. Early designs should be made visible to stakeholders to enable feedback, validate expectations, and check that system behaviour aligns with the defined problem and risk posture. Learning with stakeholders at this point reduces the likelihood of late stage rejection or over confidence and is key as part of an agile and iterative approach to design. This is fundamental in the double diamond model.

© MASS Consultants Ltd

Deliver: building trust, integrating context and sustaining value

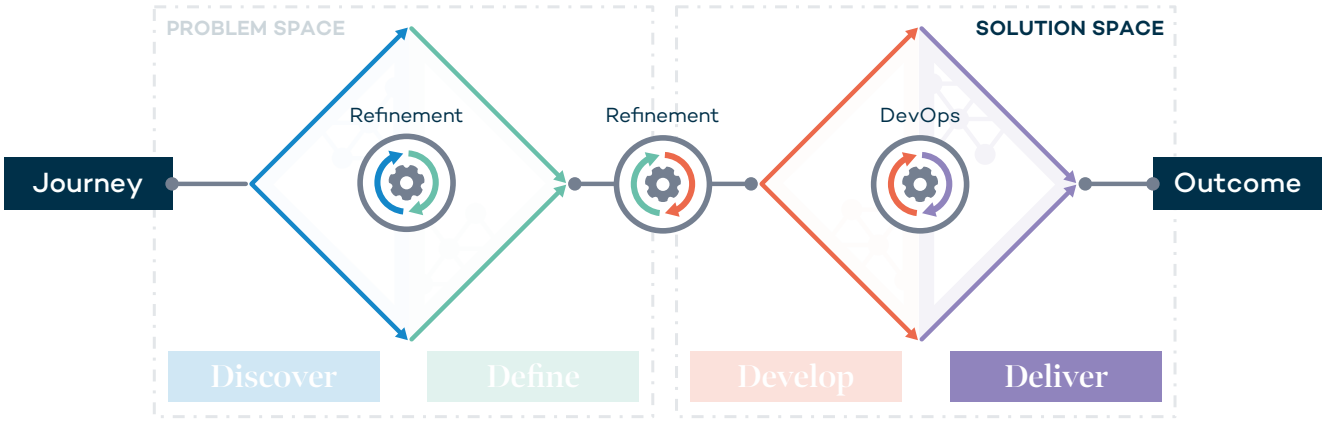


Figure 10 – The Deliver phase realises the outcomes but considers support & continual improvement.

Delivery of generative AI solutions should be approached as an ongoing process rather than a single deployment event. Agile development practices should be used to build small, testable increments that keep learning open and allow adaptation as understanding evolves.

Testing and validation should demonstrate that the system meets the requirements defined earlier, including performance against agreed benchmarks and expected behaviour in realistic operational case studies. Safeguards should be shown to trigger appropriately when uncertainty or error exceeds acceptable limits.

Integration with real systems, data, and workflows is critical, as trust is shaped by how seamlessly the solution fits into operational practice. Without this, poor integration could develop which could undermine confidence even in technically sound systems.

Training and early life support should focus on trust calibration. Users must understand what the system is for as well as what it is not for, and, how uncertainty manifests. In addition, users must understand how to avoid misuse or over reliance. Making uncertainty visible and acceptable is central to responsible use. This makes the case that users matter when it comes to designing effective and successful technology[7].

Success consideration

Success here is about ensuring that a controlled, testable and safe AI component exists, that we can learn from without adding in poor assumptions. Behaviour, risks and limitations need to be visible.

7 Oudshoorn and Pinch (2005) How Users Matter: The Co-construction of Users and Technology

© MASS Consultants Ltd

Finally, organisations should recognise that AI systems are ***never truly finished***. Continuous monitoring of behaviour, performance, and risk is required, with updates made as context, data, and usage evolve. Sustained confidence is what ultimately delivers value long-term.

Contemporary industry evidence reinforces this. The DORA 2025 research on AI-assisted software development found that AI often acts as an amplifier of existing organisational strengths and weaknesses. Below is the DORA AI Capability model, a framework for identifying seven key practises proven to amplify the positive impacts of AI on organisational outcomes.

Organisations achieve greater value when they integrate with a healthy data ecosystems, accessible organisational knowledge, strong engineering practices and user-centred ways of working[8]. This suggests that long-term AI success depends not only on the technology itself but on the organisational environment into which it is introduced.

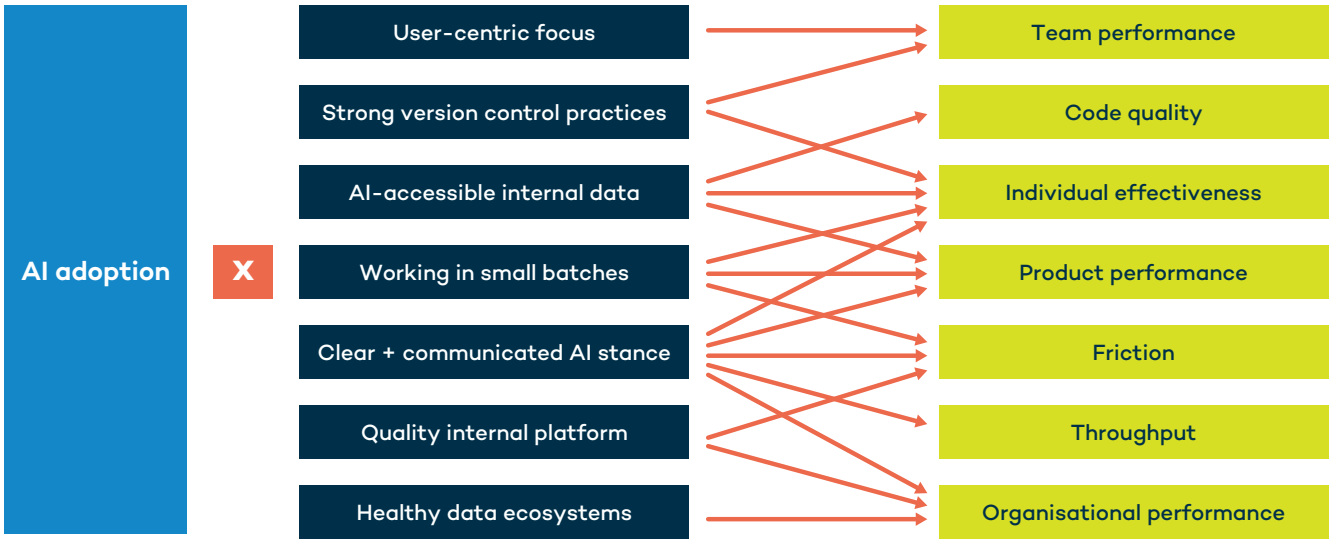


Figure 11 – The DORA AI Capability Model

8 (*DORA | DORA Research: 2025 Overview*, n.d.)

MASS
A COHORT PLC COMPANY

Cross functional framework success criteria

The framework described provides structure to delivering AI projects. To help understand specific aspects of concern and how these navigate across the double diamond model, such as elicitation of requirements, the following cross functional diagrams are provided for additional context.

REQUIREMENTS

Figure 12 looks at how project requirements are supported by the different phases of the framework. Managing requirements of an AI project can be challenging as it can be very difficult to have a binary pass/fail, as the system is non-deterministic. Our goal therefore is to focus specifically on discovering how an AI system is to be employed in the operational workflow, its responsibilities and what instructions, data and policy it needs to follow to realise its given objectives. If we understand this, we are more able to produce 'quality requirements' that are representative of the real-world operating conditions and easier to measure the system against.

The Discovery phase of the framework sets the 'ground truth' understanding of the specific problem to be solved, the operating context and where it is intended that AI will contribute, in part, to solving it. The division of human / AI roles & responsibilities are discussed along with eliciting any legal, ethical, technical, safety or risk constraints involved in that division of labour. As we do with human staff, consideration is also given to defining the levels of supervision, monitoring and audit that the AI components will require.

The Define phase develops the concise and meaningful requirements that iterates with the Discover phase until a common understanding and consensus has been reached. Functional requirements identify how the AI components operate in the workflow and the non-functional requirements set out the constraints and safeguards.

The Develop phase is concerned about human review and is the opportunity to inform potential candidate technologies, consider experimentation and hone the balance of human / AI division of responsibility. It can help to surface operator training needs and help to get the level of supervision and explainability to the right levels.

© MASS Consultants Ltd

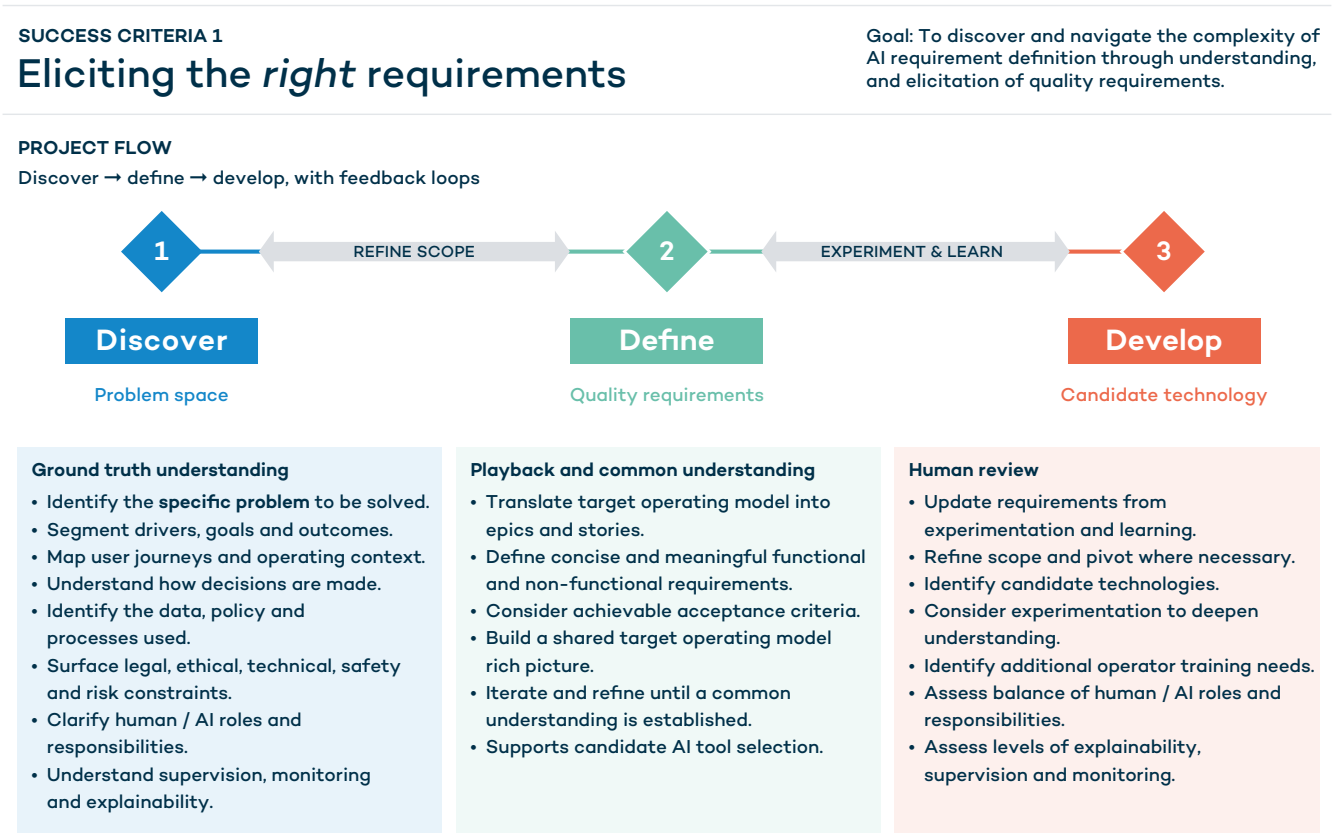


Figure 12 – How eliciting quality requirements spans Discover, Define and Develop phases.

TEST AND EVALUATION EFFICACY

Figure 13 considers how the framework supports test and evaluation efficacy. Importantly it considers developing a range of operational scenarios to ensure that the AI system can be measured against meaningful benchmarks, rather than just focussing on technical build quality. Having near real-world representations of how the AI will contribute to the operational workflow will allow visibility of expectations during testing and evaluation, leading to a better understanding of any changes that need to be made.

It uses artifacts generated in the overall Discover phase as data objects to help define the Operational Scenarios that will test the system to gauge its effectiveness. These scenarios can be assessed by human Subject Matter Experts to create a baseline result of specific expectations and behaviours, and this will provide a repeatable test to monitor any changes to the system.

During the Develop phase, areas such as assessing the adequacy of legal, ethical, technical and safety measures take place and the provision of evidence for risk owners. For example, this assessment of adequacy could highlight the need to adjust the human / AI balance where there is an ethicality concern in

© MASS Consultants Ltd

the workflow and require human judgement. Prior to test readiness reviews and formal testing, test plans are generated that meet the operational scenarios that have been developed.

The Deliver phase is where formal acceptance takes place, the system receiving early life support, staff trained and initiating a schedule of behaviour & performance monitoring. This phase is about making sure that the users of the system in operational use are:

- able to understand what the AI components are responsible for in the process.
- able to gain confidence and trust in its use.
- equipped to recognise when the system performance or expected behaviour changes.
- receiving the intended benefits and value originally planned.

And that risk owners are content with:

- the risk evidence.
- the legal, ethical, technical and safety adequacy.
- the supporting processes for monitoring and continual assessment.

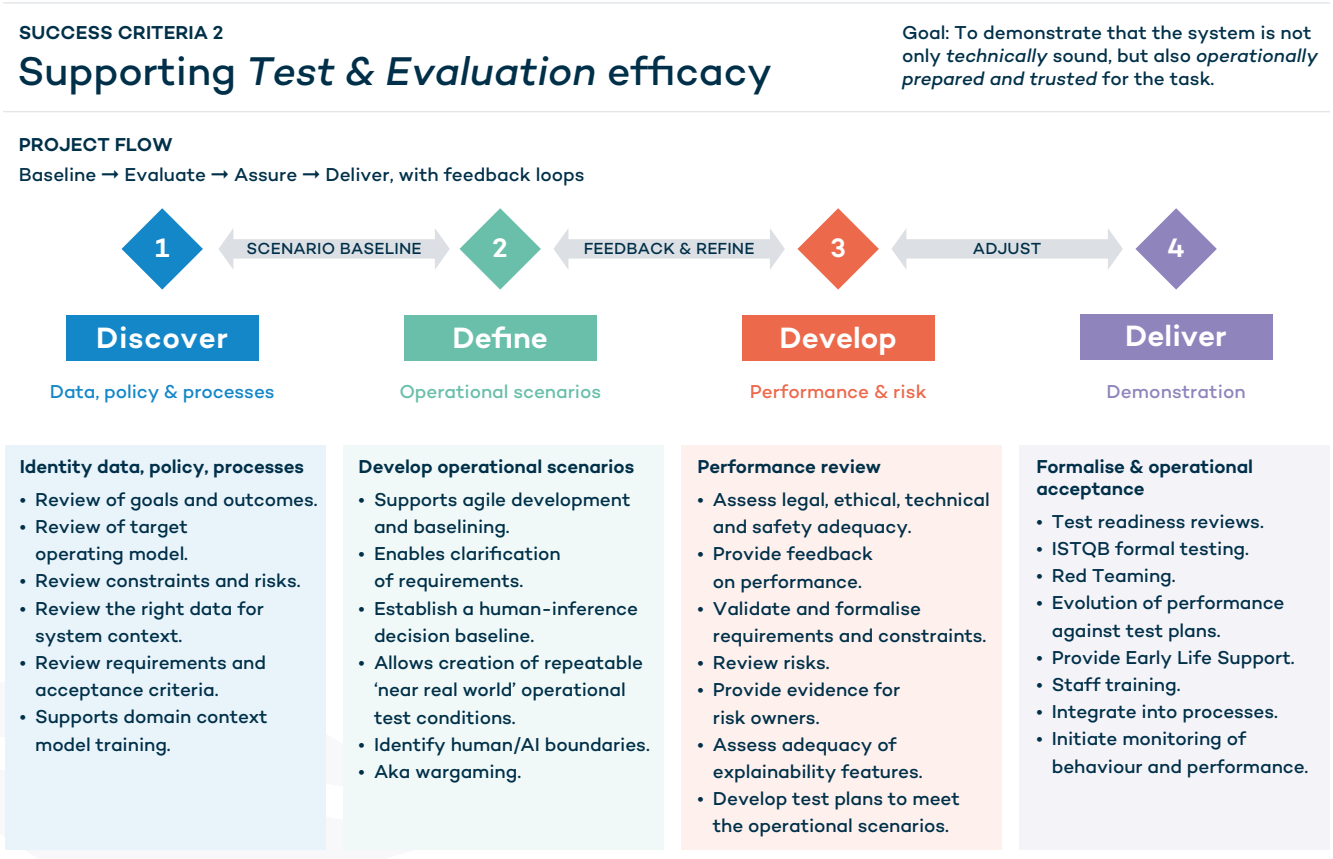


Figure 13 – Aspects of test & evaluation across the framework aligned to the operational context of how it will be employed.

ENDURING SYSTEM PERFORMANCE

As described in the Deliver: building trust, integrating context and sustaining value section, we suggested that organisations should recognise that AI systems are ***never truly finished***. As shown in Figure 14, continuous monitoring of behaviour, performance, and risk is required to ensure that the system remains operationally fit for purpose. Factors affecting a change to the system include:

- A change to the operational workflow or processes in which the system is employed.
- Changes to the data that the system uses for context, such as Standard Operating Procedures, regulations or policies.
- Updates to the technology used in the system to improve accuracy, fix bugs or improve security.
- Following a review of risks.
- On identifying a legal, ethical, safety or security issue.
- Following degraded performance or explainability issues.

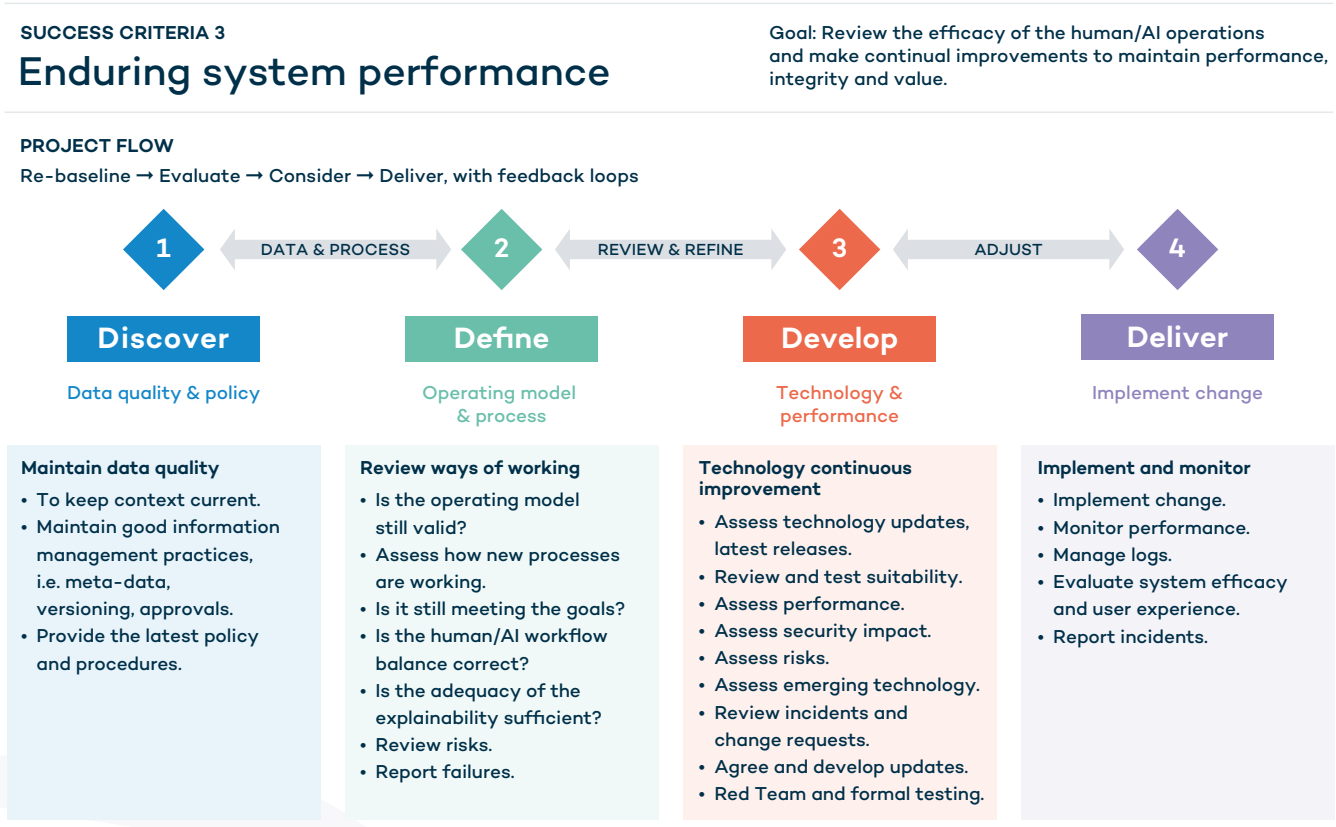


Figure 14 – A cycle of continual monitoring, evaluation and system updates is necessary to maintain operational value.

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

From Socio Technical Risk to Structured Adoption

The socio technical considerations outlined in this paper highlight a central challenge for Defence: the risks associated with Generative AI adoption do not arise solely from the technology itself, but from how that technology is introduced, understood, trusted, and used within an operational context. Technical improvements alone are insufficient to mitigate these risks, as outcomes are shaped by human behaviour, organisational norms, and the decision environments in which AI outputs are consumed.

This creates a fundamental adoption problem. If AI is introduced as a technology first – driven by capability demonstrations or performance benchmarks – limitations and unintended consequences are often discovered too late in the process, when systems are already in service and trust has been established or eroded.

In a Defence setting, this late discovery of risk can lead to rework, reduced confidence in decision-support systems, and in some cases the rejection of otherwise valuable technology. To address this challenge, AI adoption must begin upstream of technology selection. It must start with a structured understanding of the problem to be solved, the decision being supported, the operational environment in which that decision is made, and the tolerance for error associated with acting on AI-generated outputs. This requires deliberate consideration of both technical and socio technical factors as part of the problem definition, rather than treating human and organisational considerations as secondary concerns.

Design Thinking offers a practical and well-established means of achieving this. In particular, the framework at Figure 6 shows how it can provide this a structured, collaborative approach to problem exploration and solution development, ensuring that the problem space is thoroughly understood before solutions are designed and implemented. By separating discovery and definition from development and delivery, the framework supports early experimentation, validation, and learning – allowing risks to be identified and addressed when they are cheapest and safest to resolve.

When combined with a socio technical perspective, the Double Diamond framework enables Defence organisations to assess not only whether AI can be applied, but whether it should be, and under what conditions. It provides a mechanism for aligning AI capabilities with operational needs, decision criticality, and human judgement, ensuring that AI systems are designed to support – not replace – Defence decision making.

© MASS Consultants Ltd

MASS
A COHORT PLC COMPANY

Learning from AI failures: why structured adoption matters

Several high-profile failures show that AI projects rarely falter because the concept of AI is inherently flawed. More often, organisations move too quickly from a compelling demonstration to deployment without fully understanding the decision context, user behaviour, data limitations, governance constraints or tolerance for error.

A Double Diamond approach cannot guarantee success, but it can expose assumptions, risks and failure modes earlier—when they are safer and less costly to address. RAND's review of unsuccessful AI projects identifies recurring causes: misunderstanding the problem, lacking adequate data, pursuing the latest technology instead of a real user need, insufficient infrastructure and applying AI beyond its current capabilities. These causes align closely with the risks this framework is designed to address.

Failed projects are instructive because they show what happens when organisations start with a solution before understanding the problem. IBM Watson for Oncology struggled with the complexity and variability of real clinical decision-making. Amazon's recruitment tool reproduced bias embedded in historical hiring data. Google Flu Trends overestimated the reliability of search behaviour as a proxy for illness. Microsoft Tay failed when real-world user behaviour differed sharply from design assumptions. These were not simply technology failures; they were failures of discovery, definition, validation and governance.

Viewed through the lens of the Double Diamond, each example raises questions that should have been explored earlier: What decision is being supported? whose judgement remains essential? what data can be trusted? what assumptions are being made? what are the consequences of error? and how much confidence is required before deployment?

By systematically addressing these questions before implementation, organisations place themselves in a much stronger position to realise the benefits of AI while avoiding avoidable operational, ethical, and organisational risks.

© MASS Consultants Ltd

Case	What failed	What Double Diamond would force earlier
IBM Watson for Oncology	AI struggled to deliver consistent clinical value in messy real-world healthcare settings.	Discover real clinician workflows; Define evidence and risk thresholds; Develop with clinical validation before scale.
Amazon recruiting AI	Historical hiring data embedded gender bias into candidate scoring.	Discover bias in source data; Define fairness as a success criterion; Develop with bias testing and challenge.
Google Flu Trends	Search behaviour was treated as a reliable proxy for disease activity but shifted over time.	Discover behavioural context; Define limits of proxy data; Deliver with monitoring for drift.
Microsoft Tay	Public users exploited the chatbot's learning behaviour within hours.	Discover adversarial user context; Develop safeguards/red-teaming; Deliver with monitoring and intervention controls.

Table 1: AI Failures and how the Double Diamond approach could have helped.

Conclusion

Adopting AI in Defence demands far more than selecting a capable model or integrating the latest technology. Defence decisions are uniquely consequential, frequently time-pressured and shaped by complex socio-technical realities. AI can deliver meaningful efficiencies—summarising information, accelerating research, identifying patterns and analysing imagery—but only when it is deliberately aligned with the problem it is intended to support.

A consistent theme runs throughout this paper: not all decisions carry the same consequence, so AI cannot be applied uniformly. Confidence, inconsistency and non-determinism introduce risks that intensify in operational environments. Human factors—trust, interpretation, experience and tacit knowledge—are inseparable from system performance. The interaction between people, processes, policy, organisational culture and technology ultimately determines whether AI creates operational advantage or adds workload, uncertainty and risk.

The MASS AI Adoption in Defence Framework addresses this challenge by starting with the problem and operational context, not the technology. Its iterative Discover, Define, Develop and Deliver phases connect operational need with socio-technical understanding, requirements, architecture, experimentation, test and evaluation, governance, training and continual monitoring. This creates deliberate opportunities to challenge assumptions, compare AI with non-AI alternatives and redirect activity before cost and risk become embedded. It also makes assurance a lifecycle discipline rather than a final checkpoint.

© MASS Consultants Ltd

AI systems evolve, operational contexts shift, and user behaviour adapts over time. Sustainability therefore depends on enduring monitoring, training, support, and recalibration of trust. Defence must recognise that AI systems are not "finished"; rather, they are maintained, refined, and governed throughout their lifecycle. Applied in this way, AI becomes neither a replacement for Defence expertise nor a technology in search of a problem, but a governed capability that augments people, preserves accountability and earns confidence through use.

The question is no longer whether AI can help Defence; it can. The real question is where it should be applied, under what conditions and with what oversight. Purpose-designed AI—grounded in evidence, aligned with operational need and supported by human judgement—can become a powerful, trusted capability. A structured, problem-led adoption framework gives Defence the best chance of realising that value while preserving safety, accountability and operational effectiveness.

Head office:
Enterprise House | Great North Road
Little Paxton | St Neots
Cambridgeshire
PE19 6BN
t: +44 (0)1480 222600

Lincoln office:
New Horizons | Teal Park Road
Teal Park | Lincoln
Lincolnshire
LN6 3AD
t: +44 (0)1522 502050

e: contact@mass.co.uk
mass.co.uk

MASS
A COHORT PLC COMPANY

Copyright © 2026 Mass Consultants Ltd. All Rights Reserved.