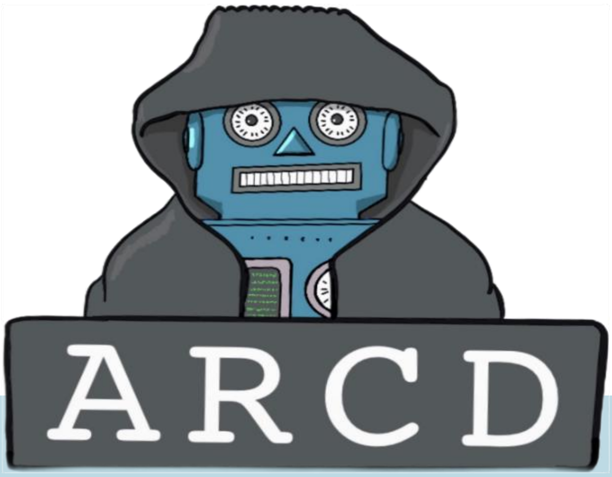
FRAZER-NASH
CONSULTANCY
A KBR COMPANY



NEWSLETTER

October 2023
ARCD Track 1: Frazer-Nash Consultancy

Approved by: Steve Little
ARCD Track 1 Programme Lead

ARCD TRACK 1 – OCTOBER 2023

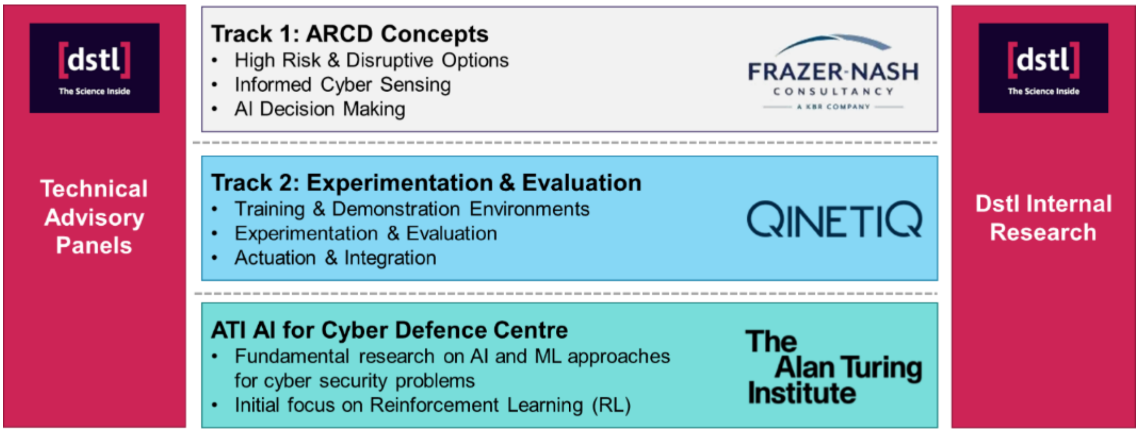
AUTONOMOUS RESILIENT CYBER DEFENCE

Due to both the increasing complexity of military networks and systems, and the sophisticated approaches of aggressors, it is becoming increasingly difficult for cyber-defenders to respond quickly and effectively to incidents. ARCD aims to mitigate these threats by compiling advanced and novel research that can be matured to be applied into defence environments.

The project objectives are to achieve the following:

- The creation of a concept demonstrator capable of autonomously responding to cyber-attacks in the context of a military environment and mission.
- Enhance cyber and Machine Learning (ML) skills in the UK supply chain.
- Further developing understanding of strengths and limitations of ML technologies and their application into Cyber Defence.

Frazer-Nash has commenced year three of the four-year ARCD research project, leading the MOD's research into Generation-After-Next (GAN) Cyber Defence concepts. ARCD is funded by the Defence Science and Technology Laboratory (Dstl) with the goal to develop self-defending, self-recovering concepts for military operational platforms and technologies and an aspiration to achieve 'Full Auto' Cyber Defence.



Technical Advisory Panels

Track 1: ARCD Concepts
- High Risk & Disruptive Options
- Informed Cyber Sensing
- AI Decision Making

Track 2: Experimentation & Evaluation
- Training & Demonstration Environments
- Experimentation & Evaluation
- Actuation & Integration

ATI AI for Cyber Defence Centre
- Fundamental research on AI and ML approaches for cyber security problems
- Initial focus on Reinforcement Learning (RL)

Dstl Internal Research

[dstl] The Science Inside
FRAZER-NASH CONSULTANCY A KBR COMPANY
QINETIQ
The Alan Turing Institute

Track 1 delivered through the Serapis Lot 6 framework by Frazer-Nash, currently run three technical streams:

U75a HIGH RISK AND DISRUPTIVE OPTIONS (HRDO) – Technical Lead: Alex Revell

To invest in High-Risk / High Reward (HRHR) transformational opportunities and bold new ideas that may have the greatest impact to the ARCD programme.

U75b INFORMED CYBER SENSING (ICS) – Technical Lead: Steve Little

To understand a threat, and the behaviour of the attacker, to enable appropriate autonomous responses to detect, defend or deter cyber adversaries.

U75c AI DECISION MAKING (AIDM) – Technical Lead: Ian Miles

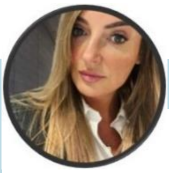
To mature concepts, introducing more complex reasoning for autonomous cyber defence response planning within representative military environments.

ARCD TRACK 1 - KEY CONTACTS

If you have specific questions related to any of the three Frazer-Nash Track 1 technical streams, please feel free to reach out of the Technical Leads or our Supplier Relationship Manager:

Steve Little – ARCD Programme Lead & ICS Technical Lead – s.little@fnc.co.uk
Ian Miles – AIDM Technical Lead – i.miles@fnc.co.uk
Alex Revell – HRDO Technical Lead – a.revell@fnc.co.uk
Gemma Whitley – ARCD Supplier Relationship Manager – g.whitley@fnc.co.uk

For any additional queries on Track 1 tasking or contracting please contact arcd@fnc.co.uk and we will get back to you as soon as possible.



SUPPLY CHAIN

Gemma Whitley – ARCD Supplier Relationship Manager

ARCD Track 1 began in May 2022 and it has gone from strength to strength since then. We have contracted 61 tasks and committed over £15m of funding which is incredible. The team have worked hard to ensure any contracts placed are done so as efficiently as possible, within the 10-day contracting period that we set ourselves as a target. Over the last five months, we have been working with Dstl to uplift the ARCD Track 1 finances which will be split across the three technical areas: HRDO, ICS and AIDM. I am pleased to announce this has now been approved, which is fantastic news for our Supply Chain. We look forward to seeing the innovative ideas you continue to have for us.

We are excited for our second Supplier "Show and Tell" Event on the 17th of October 2023 in Leeds for both Track 1 and Track 2 for which the plans are now in place. Thank you to all who have replied confirming your attendance. This is a fantastic opportunity to witness some of the exciting work being showcased that has and is being currently delivered across ARCD. It will be a great opportunity to build and enhance relationships across the Supply Chain as we continue our path, growing the UK supplier base. I would like to thank you for your continued patience over the last few months. We fully appreciate it has been busy period for everyone and there may have been some frustrations over the amount of problem books that have been published of late. I can assure you that the team have been working hard to make sure that ARCD Track 1 runs as efficiently and effectively as possible and we are delighted to have released the HRDO Problem Book 4 for Foundation Models. Submissions will be competed in a two-round process where the first round (a single page submission of the idea and ROM etc.) is submitted by the 3rd of November. Please direct any CQs to myself via the ARCD mailbox and I will respond as quickly as possible.

At the centre of our team's mission is our strong relationship with yourselves as the suppliers, and we sincerely appreciate every opportunity to connect with you. As said before, please don't hesitate to reach out directly to me as the ARCD Track 1 Supplier Relationship Manager and the team look forward to seeing you on the 17th of October in Leeds.

HIGH RISK AND DISRUPTIVE OPTIONS

Alex Revell – HRDO Technical Lead

HRDO has embraced a high-risk high-reward appetite to drive success in the development of Generation After Next (GAN) cyber defence concepts. 'Success' is defined as pushing technical boundaries, exploring challenges and learning lessons from technical failures to improve understanding of the strengths and limitations of Machine Learning (ML) technologies and their application to cyber defence. To date, the HRDO research teams have employed SOTA techniques to address the following problem areas:

Causal Inference

Pattern of Life

eXplainable AI

Knowledge Graphs

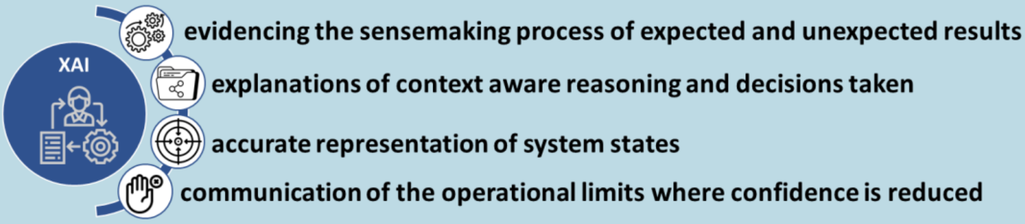
Quantum Machine Learning*

Generalisability*

Data Efficient Decision-Making

* Problem areas generated by the Open Call for Ideas Problem Book

This edition showcases some further insight into the great research that has been presented within the XAI tasks. Explainability is key to developing a user's trust, and a particularly crucial element when considering the recent progress on, and therefore challenges arising from, ML methods such as Reinforcement Learning (RL) and the ARCD context (automating responses in a fast-paced, safety critical environment). Trust is built jointly through several layers of explanation, such as:



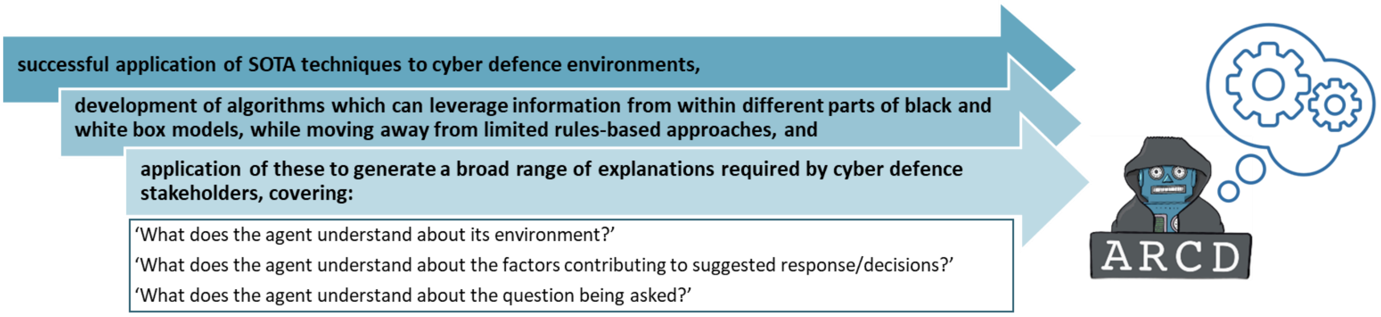
- **evidencing the sensemaking process of expected and unexpected results**
- **explanations of context aware reasoning and decisions taken**
- **accurate representation of system states**
- **communication of the operational limits where confidence is reduced**

For further information refer to the HRDO Task Summaries Section.

XAI Tasks: 10, 16, 19, 20, 22, 23, 25

Finally, providing meaningful explanations to stakeholders of the decisions underpins each of these aspects in enabling effective and efficient user comprehension during deployment, or through providing evidence and explanations during validation, verification, and assurance processes.

The research outcomes enable users to understand a surface level reason for a given action, and what an agent understands about its context. While most XAI tasks are still ongoing, the key results so far include:



successful application of SOTA techniques to cyber defence environments,

development of algorithms which can leverage information from within different parts of black and white box models, while moving away from limited rules-based approaches, and

application of these to generate a broad range of explanations required by cyber defence stakeholders, covering:

'What does the agent understand about its environment?'
'What does the agent understand about the factors contributing to suggested response/decisions?'
'What does the agent understand about the question being asked?'

Advancing the techniques and tools with which stakeholders' access explainability layers also benefits the remaining ARCD research areas. This enables researchers and developers to demonstrate intelligent, sensible behaviour or alternatively to probe more deeply into the erroneous areas of understanding explained by the XAI agents. Finally, XAI can support identification of novel cyber decision strategies, where a feature may be seen that is neither typical of a user's expectation nor evidently incorrect. HRDO's high-risk high-reward approach and the 'successes' presented here will inspire further research into the application of novel XAI techniques to cyber defence, to build trust in autonomous agents and enable them to be confidently deployed and used on real networks.

INFORMED CYBER SENSING

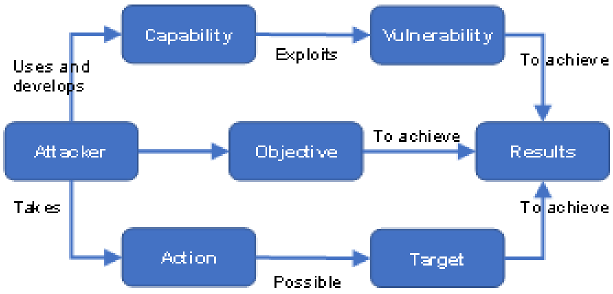
Steve Little – ICS Technical Lead

ICS is the foundation upon which HRDO and AIDM are building their agents and the team consists of myself as the Technical Lead with Holly Gregory, Charley Eccles and Claire Carter as the ICS Task PMs. We are undertaking research to develop Machine Learning (ML) techniques to provide insight into the threat and behaviour of the attacker and we are particularly focussed around the following three phases:

- **Perception** - to perceive the status and attributes of a system in the context of the environment it is operating in.
- **Comprehension** - to understand and interpret the significance of objects and events perceived within the system.
- **Prediction** - using the above reasoning to predict future events related to the system, or what steps have been taken to reach the current understanding.

As of October 2023, we have contracted 18 tasks totalling £4.4m, contributing our understanding to the situational awareness of a cyber-attack characterised against the figure to the below. To achieve this, we have built on previous work undertaken by Dstl around Serapis Lot 6 U60 and Predictive Cyber analytics.

To autonomously defend against cyber-attacks, automating Intelligence Preparation of the (cyber) Environment (IPOE) will enable the perception phase, attributing the attack will support comprehending and predicting what is happening.



IPOE – To support attribution, an understanding of the operating environment for the system is required, specifically user, traffic and network information. Work has been undertaken as part of ICS, which is contributing to automating the IPOE, this has included applying ML techniques to: identifying critical assets within a network, developing a proof of concept using Topological Data Analysis (TDA) to address what methods of transforming data are most effective for preserving cyber features, and validating the relevance of Log Source Qualification when applied to understanding cyber security.

Cyber Attribution – This is currently a human-centric approach often requiring knowledge and experience. As part of Informed Cyber Sensing we have applied Machine Learning to the problem of autonomously attributing a cyber-attack to a threat actor. As an example, this has included combining malware analysis tools, network traffic and telemetry data to generate features to be used by ML algorithms.

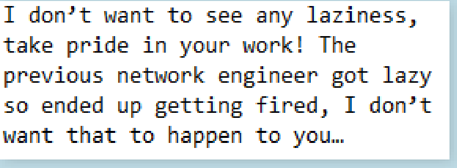
In the future we will be issuing a small number of focussed problem books, specifically looking at how we can fill gaps in our current understanding of what is required from ICS. We will also look at how we can attribute data provided from ARCD Track 2, making the outputs consistent to the STIX model we are developing for ARCD.

AI DECISION MAKING

Ian Miles – AIDM Technical Lead

With a recent flurry in contracting, AIDM has committed funding up to £5m across sixteen projects and has three more long-term contracts under negotiation. Our first 18 months has seen many landmarks, including:

- BMT (with ADSP) demonstrating the superiority of multi-agent RL over single agent in a ‘primary’ level environment, and RL agents autonomously adopting different cyber specialisms (container & eradicator). This project was briefed to the Minister for the Armed Forces at the MoD’s Defence Command Paper 2023 launch event as an exciting example of Dstl’s S&T research and industry collaboration (team pic, left).
- Cambridge Consultants standing up their sophisticated Generic Vehicle Architecture environment to train and evaluate their CO-DECYBER multi agent concept.



I don’t want to see any laziness, take pride in your work! The previous network engineer got lazy so ended up getting fired, I don’t want that to happen to you…

- Foundation of an entirely new cyber/ML company, Trustworthy AI (TAI), who have utilised GPT4.0 to build 50+ tactical network topologies to train against (somewhat disturbingly needing to threaten the LLM, see snippet left). TAI are now in negotiations for a long-term ARCD contract and are actively recruiting as a result.
- ARCD’s first example of a ML cyber defence agent outperforming a human expert in ADSP’s MVP project (with BMT), which also saw Track 1’s first integration into the PrimAITE environment. We are commencing a second MVP project exploring transfer of agents trained in PrimAITE to the high-fidelity Imaginary Yak environment, with independent evaluation by Track 2.
- On-boarding BT to expand their award winning Inflame tool to develop a more sophisticated Primary Level red agent, trained on Track 1’s largest Primary Level topology (100+ nodes).
- Xewli’s first ARCD / Serapis Lot 6 contract to explore ‘AI building AI’ starting with the Topology and Weight Evolving Artificial Neural Network (TWEANN) concept.
- Presented at two international conferences: BMT (CAMLIS) and Cambridge Consultants (SECAI).
- A notable mention should also go to Exalens, initially funded under DASA, demonstrating ARCD’s first ‘end-to-end’ autonomous cyber defence of a ‘real’ system (ROSbot) against a real cyber-attack (remote desktop protocol brute force). Follow-on to this task is being explored in AIDM.

Our outstanding progress was key in securing the significant Track 1 uplift and clearly would not have been possible without the incredible minds of our suppliers, so a massive thank you to you all. Our key focus out to March 2025 will continue to enhance the maturity of the most promising ARCD concepts, through to demonstration on a ‘representative system’. Whilst we work closely with our Dstl and Track 2 colleagues to define and utilise that representative system, current plans include scaling up to demonstration on a ‘real’ maritime platform management system and increasing use of the ARCD Imaginary Yak environment. We are also excited to soon start exploring opportunities arising from the most promising HRDO concepts and testing them in higher fidelity environments.

ARCD TRACK 2 - QINETIQ

Will Bowers – Track 2 Programme Lead

Track 2 of ARCD - Experimentation and Evaluation, led by QinetiQ, aims to determine if solutions are relevant to the unique military challenges and environments in which an ARCD capability might be deployed and whether they offer benefits over human driven responses. The scope includes the creation of information and technical artefacts that suppliers require to deliver Track 1, including training environments for cutting edge AI and ML approaches; development of a software framework to integrate the outputs of Track 1 with a set of demonstration environments; and creation of the tools and techniques to evaluate the performance and impact of Track 1 approaches.

Track 2 has similarly identified twenty distinct research questions which need to be answered if the customer is to be able to evaluate ARCD systems – both within and beyond the ARCD project itself. Within the year, the track has made progress in answering these research questions by means of (a) reports on the state of the art with respect to metrics and methods for evaluating agentic AI technologies and the cyber-defensive systems that use them; and (b) exploratory data science work, developing and testing analytics on agents trained in Task 1 environments.

In Year 3, the project plans to begin a programme of experiments, to provide additional, empirical evidence with which to inform answers to these questions. For more information on Track 2 please contact: ARCD-Track2@qinetiq.com.

THE ALAN TURING INSTITUTE

Funded by and research partners with Dstl, the ATI are leading the AI for Cyber Defence (AICD) Research Centre and providing expert insight into areas of Track 1 tasking. The centre is led by principal investigators Vasilios Mavroudis and Chris Hicks who are computer security researchers, seeking to fundamentally transform the way in which we secure digital systems through the development and application of cutting edge, deep-learning based approaches to intelligent agents. The current focus areas are as follows:

- Autonomous cyber operations and network defence
- AI for Systems Security
- Adaptive fuzzing and state-machine learning
- Cryptographic ciphers, protocols and their implementations

Whilst dedicated to solving security and privacy problems, the ATI are currently researching a variety of Deep Reinforcement Learning (DRL) techniques. DRL and related techniques offer a mechanism for planning strategically that we intend to show, through this project, can transform our understanding of, and capacity to attack and defend, computer systems and networks. For more information on the AICD Research Centre please contact: aicd@turing.ac.uk.

ARCD TASKING OVERVIEW

The following pages provide an overview of each task contracted on the ARCD programme.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 1: QTSL - Autonomous Decision Making for Cyber Defence [complete].

A 6-month project follow-on project aiming to use the Mission Planner tool to remotely conduct vulnerability scans and remediation on Virtual Machines (VMs) that were deliberately defined with vulnerabilities. The project was able to demonstrate remote vulnerability scanning, using OpenVAS, but could not complete the full remediation process within the project timescales. OpenVAS was also found to miss several of the deliberate vulnerabilities and Nessus was recommended for future work in this area.

Task 2: Camulos – Interpretable Causal Models for Cyber AI Agents [ongoing].

An 18-month research project focussed on the development of novel cyber defender RL models that display interpretable behaviour and causal explanations within a reference ‘game’ simulation environment, such that models’ behaviour and explanations can be readily understood by cyber defender operators, operational planners and decision-makers.

Task 3: University of Kent – Explainable Generic Design, Self-Evolving Intelligent Security Systems for Cyber Attack Detection [ongoing].

A 12-month study with an aim to develop next-generation intelligent security systems to monitor network activities and categorise threats in real time. The team intend to combine online clustering techniques and prototype-based approaches to construct an ML-based security system to achieve a transparent structure and explainable internal reasoning of decisions made, such that the system can continuously learn new knowledge (through discovering new prototypes representing unseen patterns) from new data but free from the problem of system obesity (namely, the system becomes oversized and less transparent due to more knowledge learned from data).

Task 4: Cambridge Consultants - Generic Persona Classification to Generate Pattern of Life [complete].

This 5-month project developed a proof-of-concept bot (PoLly) with which can generate believable and variable network traffic for use in a training or deception-based use case (i.e., green agent traffic as background noise in AI training, or honeypot). This bot has been developed to be tuneable to exhibit different behaviours from different network users through: a Characterisation work package, where 3 representative military personas were developed and combined with an application study to provide the underlying information to create the bot within the second work package; and a Modelling work package, where a Hidden Markov Model was used to generate time dependent events (based on the parameters defined by the personas) triggering network traffic generation from a set of applications.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 5: Montvieux - Data-Efficient Decision Making [complete].

A 6-month project which addressed the challenge of how to carry out Data-Efficient Decision Making by investigating the use of World Models and Model-Based Reinforcement Learning (MBRL), to:

- generate synthetic training data to effectively increase data efficiency of decision-making, and
- to reduce the amount of training data required to produce policies in cyber environments.

The authors conclude that 'Accurate world models of abstracted cyber environments [can] be trained, and [can] accelerate learning of policy optimisation methods'. Key takeaways include a method for predicting observation transitions in discrete observation spaces that have increased data efficiency in abstract cyber domains. The project has also observed that model-based approaches are limited in simple environments, and further research is needed to understand the full potential of MBRL in cyber domains. Additionally, improving data efficiency of RL algorithms helps the ARCD program by increasing the compatibility of RL to cyber domains.

Task 6: Frazer-Nash Consultancy with OxBrdgRbtx- Quantum Hybrid approach to Reinforcement Learning to train agents with less data [complete].

A 5-month project where the team challenged themselves with applying quantum technology to the cyber domain (previously unexplored in the literature) to address the problem of insufficient data available to train an autonomous RL agent classically. Their results went some way to proving their overarching hypothesis that 'QML, implemented on a D-WAVE, extracts additional insights from each training piece of data, reducing the data requirements for training such models.' The research concluded that:

- A quantum-hybrid approach can provide improved learning per datapoint, when compared to an equivalent classical approach with an identical sized network.
- The apparent improvement provided by a quantum-hybrid approach is not as significant as the improvement provided by choosing a more complex, cutting-edge algorithm solved with a much larger neural network.
- For the D-WAVE Advantage system, there is a limited increase in solving time for larger problems.

Task 7: Decision Lab with Actica Consulting - Causal Inference through Neuro-symbolic AI [complete].

In this 6-month task, the team applied Neuro-symbolic AI to demonstrate the utility of causal inference in both a classification and timeseries problem in the cyber security domain (specifically identifying and responding to malicious insider threats). Neuro-symbolic AI is designed to allow the explicit nature of symbolic AI to shine a light into the “black box” of the more traditional neural network-based machine learning through combining the two paradigms: neural networks and symbolic AI. Overall, the team implemented a Logical Neural Network (LNN) model and which resulted in some success in determining the difference between malicious and non-malicious users based on the scenario and individual user in question.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 8: Decision Lab with Actica Consulting - Data-Efficient Cyber Investigation Determination and Evaluation (DECIDE) [complete].

A 5-month Project where the team successfully demonstrated that:

- AI training data can be utilised to identify high risk components in a model network, and;
- The DECIDE tool can evaluate the quality of agent training, distinguishing between agents which successfully and unsuccessfully complete their tasks and the actions taken.

Their approach makes use of model free techniques, allowing an unbiased, confident diagnosis of capability risk profiles for prioritisation. The project has established that the DECIDE tool has the potential to complement the existing Multi-Criteria Decision Analysis (MCDA) -based approaches to risk prioritisation (highly qualitative assessments which aggregate prior subjective judgements reflected in documentary data sources).

Task 9: Decision Lab with Actica Consulting - Data Efficient Model Inference (DEMI) [complete].

A 6-month project to provide a proof of concept showing that model-based RL augmented with intrinsic motivation and curiosity agents can be trained to the same level of effectiveness, using less data than traditional model free-approaches. To tackle this, the team explored the potential of world-models, specifically leveraging the capabilities of DreamerV3, as a mechanism for augmenting data efficiency in autonomous cyber defence agents. Overall, this work shows promise that use of world models provides a more data-efficient solution for cyber defence:

- DEMI was able to train to attain similar levels of performance as the ATI implementation while using approximately 20% of the iteration.
- The DreamerV3 agent adapted and used for DEMI required no hyperparameter optimisation to attain similar levels of performance, despite not being designed for or tested previously in the cyber domain. This has positive implications in terms of transferability and generalisation of agents to diverse settings, without having to devote as much effort to tuning as is often the case with Reinforcement Learning algorithms.

Task 10: Decision Lab with Actica Consulting - White-box Models for Explainable AI [complete].

A 5-month Project addressing the lack of trust in SOTA AI techniques to allow complex, high-performance models be understood and used with confidence by non-domain experts. The team implemented a Neuro-symbolic AI model called a Logic Tensor Network (LTN), as it combines the high predictive power of neural networks with the inherent explainability of symbolic AI. Using this, the team designed and developed a high-performance, eXplainable AI (XAI) tool around a synthetic malicious insider dataset. Whilst the model and user interface (UI) was successfully developed, it exhibited poor performance when classifying malicious behaviour. This was concluded to be due to the limitations of the dataset and that to improve performance, increasingly complex data processing and modelling would be required; likely in opposition to the concept of XAI.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 11: University of Kent - Reward shaping for network defence based on reinforcement learning [ongoing].

12-months of research investigating how to improve sample efficiency through developing heuristics for intrusion detection that can be used to define and extract high-quality (dynamic or static) shaping rewards that can easily transfer between environments so they may be used as part of a high-level self-organising meta-controller able to detect new activities and guide the RL process. Reward shaping is a justified method for improving the sample efficiency of RL as it makes the learning process more informed.

Task 12: Riskaware - Causal Inference for Cyber Security (CAICS) [ongoing].

An 8-month project exploring the benefits of implementing causal graphs and embedding contextual information within RL design to improve situational awareness etc. but also to understand the chain of causation within a cyber battle within the YAWNING TITAN environment. Two of the key research questions include:

- How might causal inference be applied to adversarial, multi-agent cyber-security relevant scenarios to examine the behaviour and dynamics of agents in the presence of cyber-relevant common cofounders?
- How might we develop on this capability to integrate contextual information into the paradigm constructed, therefore creating agents which adapt their behaviour to adversarial posture and possible scenario-driven failure?

Task 13: University of Kent - Machine Learning for Causal Inference [ongoing].

A 10-month Project using observational data to learn a response surface for predicting counterfactual responses, which can then be used to estimate the causal effect. To ensure the causal relationship is learnt, and not just correlations, valid causal predictions will be made using multiple approaches (RCT, and Covariate adjustment via single and multiple models) based off ideas from the Structured Causal Model [Halpern and Pearl 2005] and the Potential Outcome Framework [Imbens and Rubin, 2015]. The key research questions include investigations to combine the ideas from the two approaches [RCT and Covariate Adjustment] to learn a shared representation of the covariates but still have two models learning the effect of the treatment separately; and how will such an approach work for a cybersecurity setting?

Task 14: Cambridge Consultants - Mapping Cyber Causes and Effects [complete].

A 6-month project where the team explored construction of a general-purpose graph-based framework for analysing and modelling security in cyber-physical systems with the overall objective to provide greater depth and breadth of attributes and attacks, whilst remaining sufficiently abstract and flexible to enable application to a wide range of scenarios. While this framework requires a relatively rare mix of skills and experience to create and update, it is key to modelling and simulating realistic, complex, cyber-physical systems, their users, attacks and defences. It thereby allows autonomous agents to evaluate the best response and recovery methods to future cyber-attacks which will operate at machine speed and require machine-speed defensive responses.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 15: Frazer-Nash Consultancy – Knowledge Graphs [complete].

A 5-month research task originally aiming to develop an ontology and conduct a feasibility study into completion of information or deconfliction of states in a knowledge Graph (KG), which is able to explain real time activity through probabilistic output and suggestions of countermeasure, and evolving ontologies in the event of unseen information. The findings initiated a change to the original scope to develop a full implementation of problem-specific Bayesian Networks (BNs) for relevant problems seen in the cyber domain. In summary, considering a network traffic dataset and attack scenarios, the following three separate BNs were constructed to:

- Resolve conflicting or missing information through predicting properties in nodes,
- Predict the next action by a suspicious user, and
- Predict the malicious score of next attack by matching to an established attack pattern.

Task 16: BMT with Mind Foundry – XAI [ongoing].

5-months of research investigating how to develop and train an eXplainable RL (XRL) agent where explanations which fall under three different categories: Feature Importance; Learning process and Markov Decision Process and Policy-Level; and, how to best visualise the explanations made by the RL agent. This team are determining how these approaches apply to, and best present, the ‘what’ (happened and has been done), ‘when’, ‘where’, ‘why’ (have these actions been taken) and ‘how’ aspects of an explanation. The results of their initial literature review consider how a problem should split along the algorithmic and usability lines to bridge the gap between abstract mathematical outputs (targeting ML researchers) and end-user interfaces demonstrating typically bespoke visualisations (targeting cyber analysts). The output will be a ‘dashboard’ of the minimum viable visualisations required to understand which features in the observation space an agent is paying attention to when choosing its action. Visualisations will enable the user to gain a sense of scenario history and agent actions on nodal diagrams.

Task 17: BMT with ADSP - Data Efficient RL [ongoing].

A 5-month project undertaking exploration of Meta-RL to build agents within a Multi Agent Reinforcement Learning (MARL) environment that can quickly learn new tasks by leveraging prior experience on related tasks as opposed to re-training from scratch. While the team are preparing for the final few weeks of results collection and consolidation, early findings are starting to indicate that vanillia PPO shows some resilience to scenario switching and SPR provides some evidence for yielding a better worst-case guarantee while showing a longer time-to-optimality on the new scenario. *Please note these results are yet to be confirmed and should not be quoted out of context.*

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 18: ADSP - (A Generalist RL Agent for Cyber Defence) [ongoing].

A 5-month project adapting DeepMind's ground-breaking Gato agent to the cyber security domain. This team's objective is to create a single, versatile defender capable of effectively operating across multiple cyber environments, irrespective of differences in observation and action spaces. The team ultimately wish to determine whether: Can Gato be trained to reach expert performance on each environment using the same set of weights and model architecture? And what is the performance of Gato after the improvements have been made – how close is it to expert performance?

Task 19: Intellium AI - Inversion Resistant XAI [ongoing].

A 6-month project aiming to address the problem of model inversion attacks (reverse engineering the model to infer and reconstruct sensitive training data) on autonomous AI security systems building a solution that autonomously recovers while preserving usefulness/security of an XAI model. The team will investigate the concept that in many cases insights from training data can help a malicious actor target weaknesses in cyber security systems. Existing work primarily focuses on the Computer Vision (CV) space however, this project aims to re-purpose and build upon these existing works to produce inversion response and recover mechanisms for XAI models that work with network data as opposed to image data.

Task 20: Intellium AI - Mitigating Adversarial Attacks using Explainable Denoising Autoencoders [ongoing].

A 6-month project aiming to explore the use of Denoising Autoencoders and XAI techniques for response and recovery from attacks on AI systems involving manipulation of input data. The goal of the project is the design and implementation of an AI system that can detect & respond to adversarial attack and assessing its performance. Adversarial attacks on autonomous AI systems involve the deliberate manipulation of inputs to an AI system to cause it to make incorrect decisions. The attackers can use various techniques such as adding noise to the input features or generating adversarial examples to mislead the AI system.

Task 21: Intellium AI - Reliable Response Through Federated Learning [ongoing].

A 6-month project collaborating with the centre for cyber security research, to explore and address the issues regarding using Federated Learning (FL) and its susceptibility to data poisoning attacks. The project looks to apply secure aggregation and differential privacy combined with self-training methods to allow for automated recovery from data poisoning attacks.

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 22: Cambridge Consultants - XAI using Policy Dissection [ongoing].

A 4-month research task to explore Whether policy dissection can be integrated within cyber environments to analyse neural network activity and thereby identify patterns of behaviour. This research team have successfully applied Policy Dissection to the cyber domain and are continuing to investigate how this should be used to provide explanations. Where this technique originally aligns the intermediate representation of the learned neural controller with the kinematic attributes of the agent behaviour (i.e. align neuron activation in the network with the corresponding agent response so that stimulation of certain neurons can produce a desired action), this team have used this technique to determine input and output mappings which subsequently enable explanations to be provided when these mappings are re-seen.

Task 23: Frazer-Nash Consultancy with Howard Science - Genetic Grammar for Explainability Wrappers [ongoing].

A 6-month research task which tackles the lack of explainability to NN's by exploiting the encapsulation and modularisation processes in Genetic algorithms to provide explanation layers (I.e., a black box approach). The team aim to solve this by informing the production of a problem grammar using methods such as Genetic programming to reduce the complexity of the end model without sacrificing reproducibility. This enables modularization but also has the benefit over LIME / SHAP by understanding how a combination of inputs influence model predictions and explain why there is a sudden change in behaviour between areas. The key research question for this project is: "Can a powerful extension of genetic algorithm methodology provide a useful explain ability layer for a complex black box model, across a wide spread of the input space? Can the limitations of the layer be adequately explained?"

Task 24: Advai – A Purposefully Uncertain Recommender Agent for Cyber Defence [ongoing].

6-months of research to explore a new theoretical paradigm for RL called "Possibility Theory", which should allow uncertainty to be built into RL agents in a more computationally manageable way than traditional "Probability Theory". By making uncertainty estimation an inherent part of Cyber Defence, the agent will only give confident recommendations in situations that have been well explored; in unusual or rare situations that have not been seen before, any recommendations will be made with very low confidence. In short, the Advai team will demonstrate that the agent will be able to know when it doesn't know…

U75A HIGH RISK AND DISRUPTIVE OPTIONS

Task 25: BAE - XAI Model Architecture [ongoing].

A 6-month project undertaking research into XRL methods and Theory of Mind (ToM) approaches to subsequently implement a novel explicit belief model. The ultimate aim is to evaluate the accuracy of predicting hosts that a Red will target alongside future actions and type of red agent Blue is being confronted. The model should:

- Allow humans to better understand the decisions made by cyber security agents.
- Improve the decision making of cyber security agents through granting them the ability to reason over the beliefs, goals and desires of Red RL agents.

Task 26: Frazer-Nash Consultancy with Oxbrdgrbtx – Quantum Extension [ongoing].

A 3-month extension to Task 6, focussing on 3 primary goals:

- Fuller training of the 32 node DBM to allow for better comparison (Only 65K~ steps done).
- Optimisation of D-WAVE through parallelisation to reduce QPU cost.
- Test a range of larger networks (was shown that they could be added to D-wave with little cost).

The team plan to evaluate computational time and accuracy of quantum vs classical approaches on different training lengths and network sizes allowing an extended firm conclusion to the original hypothesis of ‘QML, implemented on a D-WAVE, extracts additional insights from each training piece of data, reducing the data requirements for training such models’.

Task 27: Frazer-Nash Consultancy with Oxbrdgrbtx – A Novel Application of Quantum Machine Learning to Maximise the Value of Training Data (PPO application) [ongoing].

A 5-month project, inspired by the findings presented in Task 6 and the current literature position on classical approaches (where Q-Learning itself can be outperformed by a more advanced Proximal Policy Optimisation (PPO) RL method). This research aims to combine the observed benefits of Quantum Machine Learning (QML) with a state-of-the-art Proximal Policy Optimisation (PPO) approach. The team are implementing the approach on a D-Wave quantum device and testing its performance using cyber representative datasets obtained from the PrimAITE environment. The overarching hypothesis for this work is: Does a DBM integrated into a PPO RL algorithm and implemented onto a D-wave device, show faster learning data point than a classically trained equivalent? While it is clear that QML is a novel field and currently there is very little published research on using it in combination with PPO, none of which considers using a D-Wave device, if the previous QML success with Q-Learning can be translated to the PPO algorithm, then a strong route to the exploitation of QML will be realised.

U75B INFORMED CYBER SENSING

Task 1: Riskaware – CACTI [complete].

The Critical Asset Cyber Terrain Identification (CACTI) project undertook a data-science led investigation of the applicability of selected techniques to identify critical assets within a network. This covered the use of GNNs and related technologies such as Relational-GNNs and Graph Attention Networks (GANs), to develop a methodology to create a prototype tool to address the requirement for identification of critical devices within a newly encountered network.

This work is being extended through Tasks 16 and 19 to provide a dynamic view on support of the Intelligence Preparation of the (cyber) environment.

Task 2: Frazer-Nash Consultancy - Autonomous Cyber Attribution Follow-On Work – U60 Extension [complete].

Serapis Lot 6 U60 applied machine learning to the problem of autonomously attributing a cyber-attack to a threat actor using machine learning. Frazer-Nash's work focussed on using malware analysis tools to generate features that can be used by machine learning algorithms. This task provided an improved understanding of the machine learning models that are applicable to automated cyber attribution and was combined with U75b Task 5 to understand how information from multiple sources, such as malware binaries, network traffic and telemetry, can be used to provide broader coverage of a cyber-attack.

Tasks 2 and 5 were integrated into a Proof-of-Concept JAM (Joint Attribution Model) framework, we are planning follow on work to mature JAM to provide a coherent view, in attributing a cyber-attack.

Task 3: Frazer-Nash Consultancy – Exploring Behavioural Contexts of Cyber-Attacks [complete].

This project approached the problem of cyber attribution from understanding the behaviour of the attacker, and how attacker behaviours change in relation to the victims of the cyber-attack. This was split into two phases. Phase 1 was a literature review of historic cyber-attack case studies, each involving attackers with different motivations, identifying key features of targeted organisations using the COM-B model of behaviour to create attacker 'profiles'. Phase 2 identified and compared different methods used to analyse crime data, including directional acyclic graphs (DAGs).

Tasks 3 and 4 explored the human side of attribution, we are considering contracting a follow-on activity to develop an ontology to realise the benefits of human machine teaming to support the attribution of a cyber-attack.

Task 4: Frazer-Nash Consultancy – Research into Behavioural Cyber Attribution Frameworks [complete].

Serapis Lot 6 Task U60 recommended the need for more research into how cyber attribution models contribute to cyber attribution. This work was split into two phases split into two phases: a literature review into current cyber attribution models and behavioural factors; and in-depth comparative analysis of identified models and their relevant attributional factors, recommending how different stages of threat intelligence could be automated by adding a layer of behavioural factors over existing cyber threat intelligence systems.

U75B INFORMED CYBER SENSING

Task 5: Montvieux – An attribution framework integrating malware, networking, and telemetry data [complete].

Initial work on Serapis Lot 6 U60 showed that multiple attribution facets (ATT&CK class, nation state, and APT) can be inferred from labelled malware samples using ML techniques. Network traffic and telemetry data was also identified as a potentially rich source of relevant detections. Task 5 trained new baseline models to handle networking and telemetry data and integrated this work with that from Task 2 into a 'Joint Attribution Model (JAM)' framework.

Tasks 2 and 5 were integrated into a Proof-of-Concept JAM (Joint Attribution Model) framework, we are planning follow on work to mature JAM to provide a coherent view, in attributing a cyber-attack.

Task 6: Montvieux - Cyber attribution fingerprint classification via deep learning [complete].

This task applied techniques from the field of digital imagery forensics to machine speed cyber threat detection and attribution. Creating a proof-of-concept attack fingerprinting classification tool by extracting uni-dimensional time series from network logging data and network packet captures, encoding as multi-channel two-dimensional imagery field datasets with state-of-the-art (SOTA) computer-vision based classifiers to achieve classification on attack typologies.

It is intended that future work will integrate the outputs from Task 6 into the JAM framework.

Task 7: University of Liverpool – Hidden Network Model: Learning and Attacker Identification [ongoing].

Starting with the premise that behaviour displayed by malicious agents differs measurably from each other. UoL will develop novel methods that are able to estimate the probability that an agent is responsible for a given identified attack. This task comprises of two distinct steps, initially learning a formal model of malicious network behaviour for each potential attacker as a reference point for his/her future attack behaviour, this work generalise their learning algorithm for the construction of Hidden 1-Counter Markov Models (H1MM). Step 2 provides an evaluation as to what extent the observed behaviour during exploitation of the network complies with any of these learned models. To deliver this task UoL have recruited an experienced postdoc with expertise in probabilistic models, learning algorithms and implementation of proof-of-concept tools.

Task 8: Raytheon UK - Topological Data Analysis for Network Data [ongoing].

Due to the complexity of network data and malicious actors obfuscating their activity network traffic monitoring and intrusion detection are challenging problems. Task 8 developed a proof of concept using Topological Data Analysis (TDA) to summarise data while preserving the high level "shape". TDA's have been shown offer advantages in anomaly detection, activity clustering and as the basis for activity prediction using network logging data. Task 8 investigated using Machine Learning (ML) models with persistence images to gain further insights such as classifying anomalies. The research showed some really promising initial results and we are waiting to receive suggestions from Raytheon for follow-on work to understand how to implement TDAs into the ARCD demonstrator.

U75B INFORMED CYBER SENSING

Task 9: Montvieux - Live Validation of Prototype Cyber Attribution [complete].

There is a portfolio of Open Source and Machine Learning (ML) tools which are able to observe and understand different aspects of threat actions, specifically understanding Tactics, Techniques and Procedures (TTP) being employed by adversaries, and potentially who they are. Building on Task 5 this task implemented an instrumented and online sandbox environment to support more representative tool testing (and refinement). The sandbox was opened to the internet as a honeypot collecting 'real' data. We have collected three different datasets, which we have obfuscated to remove any personnel information, and which are available for subsequent tasks.

Task 10: Montvieux – Bayesian Networks for Cyber Threat Attribution [ongoing].

Machine Learning (ML) and Deep Learning (DL) models are inherently static and deterministic once trained, this is at odds with cyber-attacks and the adversaries who perpetrate them. Bayesian models are capable of systematically dealing with uncertainty and incomplete observations. This task is investigating the application of Bayesian Networks for incorporating noisy and partial observations, as well as human expert prior knowledge, into causal reasoning models of Advanced Persistent Threats (APTs). This task will build a proof-of-concept Bayesian Network model (or models) for inferring adversary Tools, Techniques and Procedures (TTP).

We anticipate that Task 10 (along with Task 11) will be integrated into the JAM framework to support attribution of a cyber-attack.

Task 11: Montvieux – Autonomous Cyber Attribution Model Advancement [ongoing].

Typically, cyber attribution is used to infer techniques and strategies being employed, the goals motivating an attack and who is responsible. This work is exploring how to uplift technique inference to APT inference via the associations contained within the ATT&CK knowledge base, by generating a dataset linking TTPs to APT Groups, train a Machine Learning (ML) model on this data to infer a probability distribution over APT Groups.

We anticipate that Task 11 (along with Task 10) will be integrated into the JAM framework to support attribution of a cyber-attack.

Task 12: Elemendar – Active Learning for continuously improving extraction of CTI entities [ongoing].

Elemendar's READ tool uses Named Entity Recognition models to process human authored, unstructured (Cyber Threat Intelligence) CTI reports into structured CTI data. This task extended the READ system to detect model staleness or concept drift in the models, enabling analysts to configure custom rules to refine the model training process. The system will be able to select samples from the user-generated data for use in model training based on model behaviour, analyst actions and interests.

The research made enhancements to the READ, including custom domain-specific overrides to entity extraction, however we were not able to automatically select data from unstructured CTI reports.

U75B INFORMED CYBER SENSING

Task 13: Accenture – Log Source Qualification [complete].

Log sources are essential to provide environmental visibility to detect notable security events, amongst benign indicators and false positives, however these sources need to be assessed and categorised to determine the data quality and relevance to security monitoring. This task comprised of three phases: Analyse and categorise detection logic and machine data; understand the relationships that exist between data logs and cyber-attacks (mapped to the MITRE ATT&CK framework); and incorporate findings to dramatically improve the ability to detect, understand and remediate cyber-attacks. The research demonstrated the use of detection logic against machine data to derive cyber activity on a network. We are waiting for the supplier to propose follow-on work to address identified challenges and limitations, including performance issues.

Task 14: Frazer-Nash Consultancy – File System Analysis for Automated Digital Forensics [ongoing].

This task is investigating how a large set of file system metadata can be used as a knowledgebase, to classify ingress data from a file system, as suspicious, attribute it to a set of similar activity within the system and use this to understand which techniques an adversary is currently employing to attack a system. This is being achieved by taking snapshots of the file system and then use machine learning (ML) techniques to process the machine-readable data at machine speeds, potentially enabling trial obfuscation and reconnaissance activities to be identified. This is the only task which is looking to exploit file system metadata, which we will be looking to pivot to different operating systems.

Task 15: Tulpa – Modelling Adversarial Behaviour to Enable AI Predictions Analogous to Counterfactual Reasoning [ongoing].

This task is developing a prototype knowledge graph, based on causal models, that encodes the adversarial behaviour of human experts. How causal knowledge graphs might be best combined with multiple ML techniques (inverse reinforcement learning, deep reinforcement learning, multi-agent co-training and transfer learning) to train explainable/scalable defensive agents capable of mimicking counterfactual reasoning will be investigated. This should allow autonomous agents to be trained to undertake actions like automated vulnerability detection and adversarial prediction.
The research is establishing techniques to explore, visualise and explain agent behaviour and to compare it with human derived Structural Causal Models.

Task 16: Riskaware – Critical Asset Cyber Terrain Identification (CACTI) Phase 2 [ongoing].

Previous work on the CACTI project (Task 1) explored different GNN architectures and data enrichment strategies resulting in a demonstrable capability (Technology Readiness Level (TRL 2)) this showed significant promise for identifying critical assets. Introspection analysis was also conducted to show how the model performed when treating networks as time-series, providing insight into criticality probabilities over time. This second phase of work aims to harness the current CACTI findings to further develop a capability that can, in future phases, be integrated into autonomous agents, and is designed to respond to potential cyber-attacks within a dynamic environment in which criticality can change over time. Under task 16 we are currently training and evaluating Long Short-Term Memory models as part of CACTI.

U75B INFORMED CYBER SENSING

Task 17: TBC

Not yet contracted.

Task 18: BMT – Autonomous Resilient Cyber Defence – Informed Cyber Sensing (Prediction, Deception) [ongoing].

Where a cyber-attack cannot be prevented, understanding the timely prediction of the past and future path of the attack is crucial to understand, potentially enabling attackers to deploy deception tools. This task addresses the challenges of advancing prediction and deception techniques through the development of three proof-of-concepts (POC).POC1 – Prediction Module to analyse attack data to understand and predict the attack path and potential targets; POC2 – Deception Engine generating a Dynamic Deception Environment (DDE) to monitor and understanding an attacker's behaviour, continually adapting to an attacker's interests and activities to maintain engagement and allow collection of relevant data; and POC3, integrating POC1 and POC 2 to further improve modelling of attacker's behaviour, prediction and effectiveness of deception, this enabling a better defensive response.

Task 19: Riskaware – ICS CyberAware Subsystem [ongoing].

This project is developing a Cyber Enrichment Engine integrating three Riskaware capabilities suitable for enriching agent data that can harness different measures of cyber risk on monitored environment: CyberAware Predict (developed as part of Dstl Predictive Cyber Analytics (PCA) Programme), CyberAware Resilience and Critical Asset Cyber Terrain Identification (CACTI). The Cyber Enrichment Engine uses the Structured Threat Information Expression (STIX) object notation for all Application Programming Interface (API) interactions to post ICS data to the service and by logging internal communication between components.

The key outcome of this project will be the development of the Cyber Enrichment Engine to offer data enrichment services in support of cyber defence.

U75C AI DECISION MAKING

Task 1: Cambridge Consultants – CO-DECYBER [ongoing].

3-year follow-on project applying a Multi-Agent Reinforcement Learning (MARL) approach using Deep Q-Networks (DQNs) for cyber defence of a 'platooning' scenario with leader-follower logistics vehicles. Bespoke training environment are being built based on NATO Generic Vehicle Architecture (GVA) utilising a simplified model of the Distributed Data Service (DDS) protocol. The project also includes development of a physical demonstrator to explore the challenges of real-world edge deployment.

This project was presented at the SECAI Conference in The Hague, and a potential Army sponsor has been identified to support exploitation.

Task 2: Aleph Insights with the University of Liverpool – MIDGARD [complete].

9-month follow-on project using Gaussian Processes and Bayesian Optimisation for cyber defence in an air defence radar operator scenario. The research explored cyber defence decision making in a simulated "world" game, where the MIDGARD agent is defending the network (modelled in Yawning Titan) whilst a radar operator 'player' deploys air defence assets using sparse data in a noisy context.

A small addendum to this project is described under U75c Task 17 and follow-on proposals are under evaluation. The funding for this project included a PhD with the University of Liverpool ("RL for Physically-Aware Cyber Defence"), which started in October 2022.

Task 3: BMT with ADSP – MARL for Operational Technology (OT) [complete].

11-month follow-on project expanding to a MARL approach using Multi Agent Proximal Policy Optimisation (MAPPO) for cyber defence of an abstract maritime Integrated Platform Management System (IPMS). Key findings included multi-agent approaches outperforming single agent, agents independently taking on cyber defence roles (e.g. container, eradicator) and agents successfully defending a network when presented with partially complete detection alert data.

This project has been accepted for presentation at the Conference on Applied Machine Learning for Information Security (CAMLIS) in Arlington, Virginia. A long-term follow-on project focussed on exploitation is detailed at U75c Task 15.

Task 4: Illumr - Genetic Optimisation for RL [complete].

4-month project following-on from previous work under SECTORED PBS159. Train and test 4 SOTA RL algorithms (DDQN, DQN, PPO and A2C) using both gradient descent and evolutionary / genetic optimization in the CAGE2 Challenge. Compare performance of Convolutional NN (CNN), and Recurrent NNs (Gate Recurrent Unit (GRU) and Long Short-Term Memory (LSTM). Genetic optimisation routines trained faster than traditional gradient descent, with a slight general improvement in performance. This research was extended to more complex NN architectures under Task 9.

U75C AI DECISION MAKING

Task 5: BMT – Managing High Dimensionality in Cyber Defence Decision Making [complete].

4-month literature review exploring state of the art approaches to dimensionality reduction, in support of the ARCD goal of demonstrating concepts on a high-fidelity representative environment. The review covers supervised and unsupervised learning (primarily split into feature selection and feature extraction methods), and RL. Cyber-specific analysis was conducted to identify and characterise high value data sources for cyber defence decision-making, to support AI specialists lacking cyber expertise.

Findings are available in full report or Aide Memoire format and findings have been re-used by several suppliers via GFX.

Task 6: BAE DI – Deep RL for Autonomous Cyber Operations (ACO): A Survey [complete].

5-month study, surveying relevant DRL literature and conceptualizing an idealised ACO-DRL agent. The report provides i.) A summary of the ACO domain properties; ii.) An overview of benchmarking environments with comparable properties to ACO; iii.) An overview of approaches for scaling DRL to domains that confront learners with the curse of dimensionality; and, iv.) A survey and critique of current methods for limiting the exploitability of agents within adversarial environments.

Follow-on to this task exploring blue agent exploitability is described at Task 16.

Task 7: University of Kent with University of Newcastle – Playing Cyber Games [ongoing].

9-month project exploring the use of cyber-attack symptoms for cyber decision making, rather than causes. An initial literature developed a symptoms taxonomy aiming to provide coverage across relevant Mitre ATT@CK TTPs. The project is heavily adapting the PrimAITE environment to implement an RL Red Agent to enable an adversarial learning approach and add symptoms to the observation space.

Task 8: ADSP - Minimum Viable Product (MVP) Agent Integration [complete].

3-month fast paced project to implement the first integration of a Track 1 agent (a basic out of the box PPO agent) into a Track 2 environment (PrimAITE V1.0), followed by stress-testing of the environment. Integration was successful and environment modification recommendations were raised, which are being implemented in PrimAITE V2.0.

This project included the first known ARCD demonstration of an ML cyber defender out-performing a human analyst. A second MVP task will kick-off in October 2023 exploring training in a significantly enhanced version of PrimAITE (including multi-agent), transfer learning into a higher fidelity environment (Imaginary Yak), the first experimentation in the ARCD demonstration environment, and the first formal evaluation by Track 2.

U75C AI DECISION MAKING

Task 9: Illumr - Genetic Optimisation for RL [complete].

6-month follow-on to U75c Task 4, expanding on genetic optimization research to more complex CNN and RNN architectures within the CAGE2 environment. The project also explored transfer learning, finding DDQN approaches to be more robust than PPO in defending against new, previously unseen, enemies. Follow-on proposals are currently under evaluation.

Task 10: Trustworthy AI - Generalised Cyber Defence of Military Tactical Networks [ongoing].

6-month project researching generalisable blue agents utilising deep RL and Graph Neural Networks. A new environment framework was developed to model tactical networks, with defensive actions including decoys. Agents were trained using adversarial learning in numerous network topologies generated from GPT4 prompts in collaboration with Dstl.

This projected enabled formation of Trustworthy AI as an organisation, and a long-term follow-on project is under negotiation encouraging further recruitment in ML/Cyber skills. Defence Digital will be supporting future work by providing information on MOD's cyber infrastructure and sensors.

Task 11: BT – Automation of Response with RL [complete].

4.5-month project aiming to i) improve the level of 'primary' level offensive tooling to train blue agents, and ii) the identification of new, unseen threat pathways in a known environment. Building on their epidemiological modelling tool Inflame, BT utilised their in-house defensive cyber teams to develop RL-based red agents as more sophisticated adversaries to train against than is currently available in ARCD training environments.

Software outputs have been provided with a BSD-3 license, enabling wider adaptation and re-use within Defence. The project attracted a lot of stakeholder interest at the recent ARCD demonstration event, and long-term opportunities include working on 'real' data & systems in a sandboxed area at Corsham.

Task 12:

On hold.

Task 13: Smith Institute – Bayesian Games for Decentralised Multi-Agent Decision Making [ongoing].

6-month project in response to the 'Diversification of AI approaches' Problem Book. The project will develop and investigate the use of Adaptive Social Learning (ASL) to train decentralised social-learnt agents, acting locally, to defend a system globally. Agents can be taught to be sceptical of information from other agents when whole sections of the network are compromised and isolated, enabling distinct behaviour from "secure" network states.

U75C AI DECISION MAKING

Task 14: Xewli - Topology and Weight Evolving Artificial Neural Networks (TWEANN) [ongoing].

3.5-month project exploring TWEANN for automated cyber response as an alternative the more widely used RL methods. TWEANN uses genetic algorithms to evolve a population of neural networks until an optimal solution is reached ("AI designing AI"). The resultant NN architectures will be compared with those designed by humans in previous ARCD experiments.

Task 15: BMT – MARL for OT Phase 3 [ongoing].

21-month follow-on to U75c Task 3, aiming to continue to progress research of SOTA MARL techniques, initially including generalisability and action masking in a maritime IPMS environment. The environment will undergo significant modification to reach the 'Secondary' level, essentially a software simulation with little to no abstraction (i.e. TRL-5). A scoping study will also be conducted early in the project to identify and assess potential physical target systems for the next phase of demonstration (TRL-6).

This project was briefed to the Minister for the Armed Forces at the recent Defence Concept Paper Launch in Westminster. Exploitation on a 'real' IPMS system is our key future focus for this project. Candidates include the Dstl IPMS Test Rig, the Cyber Ship lab at Plymouth University, or partnering with IPMS manufacturer.

Task 16: BAE DI – Evaluating the Exploitability of Autonomous Cyber Operations Agent [ongoing].

3-month exploratory project exploring exploitability of blue agents in the CAGE2 environment. Here, exploitability is measured using an approximate best response (ABR) that a new opponent can learn against the fixed policy of the agent being evaluated (Blue). Using the computed ABR, exploitability quantifies how much a player (in this case Red) gains through unilaterally deviating to the ABR.

Task 17: Aleph Insights – MIDGARD Trial Addendum [complete].

1-month addendum task to U75c task 2, with the aim of providing additional benchmarks, and evaluate the performance of an agent that optimises a cyber score vs. a 'real-world' score.