Serapis Tasking Form

Tasking Form Part 1: *(to be completed by the Authority's Project Manager)*

To:	Lot 6 Frazer-Nash Consultancy Ltd	From:	The Authority

Any Task placed as a result of your quotation will be subject to the Terms and Conditions of Framework Agreement Number:
LOT 6 DSTL/AGR/SERAPIS/UND/01

VERSION CONTROL
1.2

REQUIREMENT			
Proposal Required by:	01/07/2022	**Task ID Number:**	U99
The Authority Project Manager:	[REDACTED	**The Authority Technical Point of Contact:**	[REDACTED]
Task Title:	Crafty Amulet Transparency System Learning for Users and Developers		
Required Start Date:	01/08/2022	**Required End Date:**	31/03/2023
Requisition No:	RQ0000010040	**Budget Range**	£90k

TASK DESCRIPTION AND SPECIFICATION	
Serapis Framework Lot	☐ Lot 1: Collect ☐ Lot 2: Space systems ☐ Lot 3: Decide ☐ Lot 4: Assured information infrastructure ☐ Lot 5: Synthetic environment and simulation ☒ Lot 6: Understand

Statement of Requirements (SOR)

Background

Artificial Intelligence (AI) relies upon large amounts of high quality training data, for example, data with accurate labels to train a supervised classification model. A static model may be trained with the data available at any point in time, however, if we are working with a growing dataset or a changing environment then we would want to make use of newly aquired data to provide a better model that reflects all the data, new as well as old. Therefore, a key area of AI research concerns how to make best use of new data to improve models dynamically, through continuous learning, and how to utilise user feedback to target model weaknesses, through active learning. Some of the associated issues have been considered in previous work, on 'learning in the field' (P2973S-R-015; report to be provided GFX). Continuous learning is the ability of a model to autonomously learn and adapt in production as new data comes in, while active learning is when a system can interact with a user to actively label datapoints that can improve the model. Both of these approaches contribute towards a dynamic learning process that does not require developer intervention.

The nature of dynamic learning raises important questions for trust and transparency. For example, how can we engender trust in a system that changes over time, where the outputs on one day may differ another day

and the behaviour is therefore unpredictable? There are different perspectives on trust of such a system, depending upon the role of the individual, for example, a developer, user, or regulator will likely have different concerns. Communicating the impact of continuous learning may support trust of a system, specifically if transparency is provided in the right way, depending upon the role of the user and the context of the situation, that reflects the impact of the learning process on the model throughout.

Our aim is to explore critical research questions about trust and transparency for continuous learning systems, related to understanding what the needs are, how they should be designed, and how we can evaluate and assess the impact. To achieve this, through running human participant studies, we require an expertimentation platform with an AI system deployed in a real world environment with access to real world users, including developers, end users, regulators and senior leaders.

The requirement for this activity is to develop software required for the research studies, not to conduct the research studies themselves.

[REDACTED]. The system captures new data, in terms of the utterances or questions from users, and labels them with the appropriate 'intent', comprising various functions required to provide a response. The system uses supervised learning to classify intent based upon the labelled utterance data.

This task will build upon our work to date, to develop the system to support continuous learning from new data, gathered from users as they use the system and make utterances/ask questions, including active learning capability to actively gather utterance and intent examples from users for vulnerable/weak classes.

Requirements

The requirement for this task is to develop the Crafty Amulet system (software provided GFX) further to improve the ability of the system to learn and improve intent classification, including the development interfaces that provide transparency for both developer and end user perspectives, supporting active learning. This activity should draw upon our work to date, including our analysis of the requirements expressed by developers and users to develop the Crafty Amulet system to support learning and intent classification (DSTL/TR139045; Report to be provided GFX)

We require:

1. An interface that allows developers to effectively monitor and configure the intent classification model as new data is collected and the system retrains. The interface should meet the needs of developers and advice on this will be provided by the Dstl Technical Partner (TP) informed by exisiting research (DSTL/TR139045; report to be provided GFX). Amongst other aspects as identified by the supplier and Dstl TP, the system should support, as a minimum:
    a. Visualisation to communicate existing intent concepts, associated functions, and question utterances, inter-relationships and potential intent conflicts / overlaps. This should allow for exploration and identification of capability gaps. The visualisation should utilise the concept lattice that underpins the intent architecture.
    b. Ability to evaluate performance when retraining classification model, including understanding strengths and weaknesses and impact of retraining with additional data and streams of new data e.g. new examples of utterances from users, as they interact with the system and provide feedback. The supplier will need to create dummy data for testing purposes and will not be provided with real user interaction data.
    c. Understanding of vulnerabilities in classification model and ability for developer to weed out poor data to strengthen classification model
    d. Capture and articulate model provenance over time, including inputs (data etc…), parameters, version, accuracy.
    e. Ability to compare different models, from a model library, to select best approach. Models will need to be developed as part of this activity, but it should be possible to use use external classification models i.e. via an API service such as OpenAI API. The supplier will need to create this library and identify open source models / APIs to provide demonstration of the capability.

2. An interface that allows analyst users to understand the classification model that has matched their intent. Some of the needs for users to understand the classification model involved in matching their intent have been identified in previous work (Paper Here and here), which should inform designs. Amongst other things, the system should provide:
   a. Ability for end users to understand class possibilities and identify appropriate matches for their intent, including the best available option given the constraints / boundaries of the system, via dialogue. This could also utilise the concept lattice, given the intent architecture and use for Formal Concept Analysis.
   b. Ability for end users to understand strengths and weaknesses of classification system and for the system to provide additional guidance when considering a weak class i.e. Conversational guidance for end users to enter data that will help strengthen classification system, mitigation of utterances that will weaken system. This could include actively seeking example data for weak classes i.e. active learning.
   c. Ability for end users to understand how system is evolving over time, including propagation of learning and influence of their interactions on learning i.e. the impact of utterance inputs on the model.

Once developed, support from the supplier will be required to deploy these new capabilities, as enhancements to the existing instantiation of the Amulet system (software provided GFX).

A microservice approach is preferred to an integrated solution, where the classification model and other processes should be accessible via an API service, consistent with the development of the system to date.

Procurement Strategy

☐ Lot Lead to recommend   ☒ Single Source / Direct Award

Pricing:

☒ Firm Pricing   ☐ Ascertained Costs*   ☐ Other*

Firm Pricing shall be in accordance with DEFCON 127 and DEFCON 643

Ascertained Costs shall be in accordance with DEFCON 653 or DEFCON 802.

*only at Authority's discretion

Task IP Conditions

Task IP Conditions (Follow the NIPPY guide to identify your information and IP requirements for each deliverable)	Summary of the Authority's rights in foreground IP (IP generated by the supplier in performance of the contract)
DEFCON 703 ☒	Vests ownership with the Authority
DEFCON 705 Full Rights ☐	Enables MOD to share in confidence as GFI or IRC under certain types of agreements. Can be shared in confidence within UK Government.
OTHER IP DEFCONS: 14* ☐, 15* ☐, 16* ☐, 90* ☐, 91* ☐, 126* ☐	Generally only suitable for deliverables at TRL 6 and above.
BESPOKE IP Clause ☐ *	Details to be added and agreed by IP Group
* Do not use without IPG advice and approval	
Please state in this text box if MOD or the customer has a requirement a) that one or more Other Government Departments is able to share confidentially with their own suppliers, b) to publish but you do	

not think there is a requirement to own or control the deliverable, or c) to share under a procurement* Memorandum of Understanding (MOU).

If any of these three issues applies, please contact IPG for advice before completing this form. *Listing research MOUs is not required, but can be a helpful courtesy to the supplier.

DELIVERABLES

[REDACTED]

DELIVERABLE: ACCEPTANCE / REJECTION CRITERIA

Unless otherwise stated below, Standard Deliverable Acceptance / Rejection applies. This is 30 business days, in accordance with DEFCON 524 Rejection, and DEFCON 525 Acceptance.

Standard Deliverable Acceptance / Rejection:-

Yes ☒ (DEFCON 524 Rejection, and DEFCON 525 Acceptance)

No ☐ (if no, please state details of applicable criteria below)

Deliverable Acceptance / Rejection Criteria:-

If there are any other specific acceptance/rejection criteria you would like to apply to any of the deliverables, please state them here.

Government Furnished Assets (GFA)

ISSUE OF EQUIPMENT/RESOURCES/INFORMATION/FACILITIES (*if not applicable, delete table and insert "None" in this text box)*

Unique Identifier/ Serial No	Description	Classification	Type	Available Date	Issued by	Return or Disposal Date	Any restrictions?
P2973S-R-015	The Future of AI for Defence: Learning in the field	[REDACTED]	Report	01/04/2022	Dstl	End of contract	
CS21-SER-CA-TD	Crafty Amulet Technical Design	[REDACTED]	Report	01/04/2022	Dstl	End of contract	
	Crafty Amulet system	[REDACTED]	Software	01/04/2022	Dstl		

QUALITY STANDARDS

☒ **ISO9001**   (Quality Management Systems)

☐ **ISO14001**   (Environment Management Systems)

☐ **ISO12207**   (Systems and software engineering — software life cycle)

☐ **TickITPlus**   (Integrated approach to software and IT development)

☐ **Other:**	(Please specify in free text below)

SECURITY CLASSIFICATION OF THE WORK
[REDACTED]

TASK CYBER RISK ASSESSMENT. *(In accordance with* DEF STAN 05-138 *and the* Risk Assessment Workflow)
[REDACTED]

ADDITIONAL TERMS AND CONDITIONS APPLICABLE TO THIS CONTRACT

Please ensure all completed forms are copied to DSTLSERAPIS@dstl.gov.uk when sending to the Lot Lead.

Tasking Form Part 2: *(To be completed by the Lot Lead)*

To:	The Authority	From:	The Lot Lead

Proposal Reference: _017113-97559L U99 Crafty Amulet System Learning - Frazer-Nash Proposal_ **(attached)**
Delivery of the requirement:
The proposal shall include, but not be limited to:
- A full technical proposal that meets the individual activities that are detailed in Statement of Requirements (Part 1 to Tasking Form).   - Breakdown of individual Deliverables, with corresponding Intellectual Property rights applied.   - Breakdown of Interim Milestone Payments, with corresponding due dates.   - A work breakdown structure/project plan with key dates and deliverables identified.   - A list of required Government Furnished Assets from the Authority, including required delivery dates.   - A clear identification of Dependencies, Assumptions, Risks and Exclusions which underpin your Technical Proposal.   - Sub-Contractors Personnel Particulars Research Worker Form and security clearances (if applicable)

PRICE BREAKDOWN
You are to use the costs detailed in Item 2 Table I in the Schedule of Requirement and at Annex E Table 2 of the Serapis Framework Agreement. Please also provide a price breakdown which should include, but is not limited to: Lot Lead Rates, Sub-contractors costs and rates, travel and subsistence. In support of your Proposal you are requested to provide clear details of all Dependencies, Assumptions, Risks and Exclusions that underpin your price.

Offer of Contract: *(to be completed and signed by the Contractor's Commercial or Contract Manager)*
Total Proposal Price in £	£89,801.73	(ex VAT)

Start Date:	01/08/2022		End Date:	15/03/2023
Lot Leads Representative	Name	[REDACTED]		
	Tel	[REDACTED]		
	Email	[REDACTED]		
	Date	04/07/2022		
Position in Company	[REDACTED]			
Signature	[REDACTED]			

Core Work – Breakdown

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Tasking Form Part 3:

To be completed by the Authority's Commercial Officer and copied to the Authority's Project Manager.

1. **Acceptance of Contract:**		
Authority's Commercial Officer	Name	[REDACTED]
	Tel	[REDACTED]
	Email	[REDACTED]
	Date	18/07/2022
Requisition Number		RQ0000010040
Contractor's Proposal Number		Serapis Task U99
Purchase Order Number		DSTL0000005530
Signature		[REDACTED]
Please Note: Task authorisation to be issued by the Authority's Commercial Officer or Contract Manager. Any work carried out prior to authorisation is at the Contractor's own risk.		