Centre for Defence Enterprise

CDE proves the value of novel, high-risk, high-potential-benefit research. We work with the broadest possible range of science and technology providers, including academia and small companies, to develop cost-effective capability advantage for UK armed forces and national security.

Themed competition: information processing and sensemaking

This CDE themed competition seeks solutions to challenges in the future information processing and sensemaking environment.

The total funding available for this competition is £600k.

Competition networking event: Thursday 22 May 2014 at Altitude 360, Milbank Tower, London

Competition close: Thursday 26 June 2014 at 5pm

CDE: www.science.mod.uk/enterprise
Dstl: www.dstl.gov.uk

SBRI Government challenges. Ideas from business. Innovative solutions.

[dstl]

Defence Science and Technology Laboratory

Ministry of Defence

Important information

Proposals for funding must be submitted by **5pm on Thursday 26 June 2014** using the CDE portal. Please mark all proposals for this themed competition with **'information processing and sensemaking'** as a prefix in the title (see 'technology challenges' section below for a description of the challenges under this competition).

Technical queries should be sent to dstlsensors@dstl.gov.uk. Please see guidance under the 'queries and help' section.

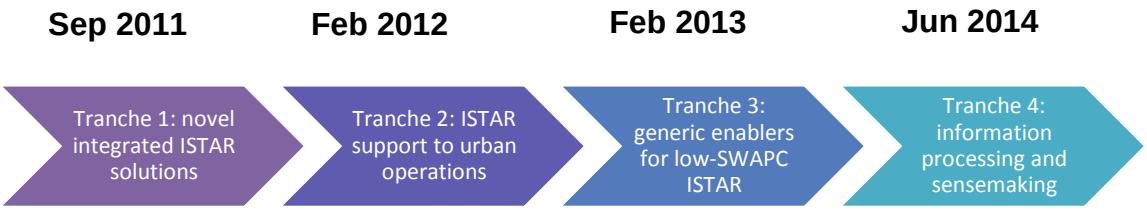
General queries (including how to use the portal) should be sent directly to CDE at cde@dstl.gov.uk.

Information processing and sensemaking

In the future[1], intelligence, surveillance and reconnaissance (ISR) is likely to be shaped by the transition "from a target-based, inductive approach to ISR, centred on processing, exploitation and dissemination, to a problem-based, deductive and anticipatory approach that focuses on ISR operations".

This means moving away from the current intelligence process, which is one of selecting targets, collecting data and processing the information/intelligence for enhanced situational awareness. Instead, we'd like to predict what might happen by using previously collected information/intelligence data and direct the collection of further information/intelligence to close gaps in knowledge.It's likely that a broader set of information sources will need to be correlated to provide enhanced situational awareness, provide indicators and warnings, and to predict and alert users of future events.

This CDE themed competition is the fourth in a series under command, control, communication and computers, intelligence, surveillance, and reconnaissance (C4ISR) concepts and solutions (CCS) project. Although the competitions are part of a series, the topics are new each time.



Series of CDE competitions under the CCS project.

Military background

Military ISR activities focus on the task of supporting informed decision making by human commanders.

In most 'traditional' intelligence-development tasks a target-based approach is appropriate. This suits well-understood, linear scenarios. For example, a target-based approach would be appropriate for the detection of hostile forces crossing a border. In this scenario, the problem is very well understood; the type of target (enemy vehicle), the physical target characteristics (size, shape, thermal and visible signature) and the area of interest (the geographic border) are all known, measurable and modelled. A military capability can be built to detect this specific target (eg a reconnaissance aircraft with a sensor of known field of view, resolution, latency, range and error). The image data collected is limited to what is needed to answer the question and this can be analysed by human operators. Also, enemy tactics are well understood (the presence of a command vehicle indicates enemy reconnaissance) and the response to this event is understood (send aircraft to attack the enemy vehicles). The process is linear, well modelled and the decision making at the centre of this target-based approach is simple.

Recent military operations have shown that the 'traditional' target-based approach is not always appropriate; (what is the 'target' in a peace support operation?) the problem, and so the task of the decision maker, is too complex.Additionally, the variables that need to be measured are often not physically definable - ideas, opinions and intentions are as important as location and number but are far more difficult to measure. Added to this difficulty, the decision maker has much more unprocessed data available that may be relevant - aerial imagery, UAVs, signals intelligence, written reports and intelligence from local civilians (each with different latency, resolution, availability, error and field of view). In this complex, non-linear scenario, more information on potential 'targets' does not always help the decision maker. Another approach is required.

Advances in communication and sensing technologies mean future decision makers will see a continued growth of information sources available, each with unique characteristics of latency, error, reliability, resolution and availability. Despite being well trained, the capacity of the future military decision maker to analyse and

[1] http://www.defensenews.com/article/20140219/DEFREG02/302190029/

make decisions in complex future operations will not match this increase in information available . A gap is likely to develop between the capacity of the decision maker to make decisions against the information available to him. A problem-based, deductive approach is therefore required to replace the traditional target-based approach.

The future technological challenge will be to remove the information overload on the decision maker whilst improving the decisions they are able to make. To do this, decision makers must be freed up from the mundane sifting, filtering and processing of raw data. They must be given useful, usable, processed intelligence in way that takes in to account the variability in quality, resolution, latency and availability of the sources without further confusion.

For further reference a detailed description of the DCPD cycle can be found at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33704/20110830JDP2003rdEDweb.pdf and examples of the military user and some types of intelligence (INT) can be found at: https://www.army.mod.uk/intelligence/32235.aspx

Technology challenges

Current government information processing and sensemaking is a human-intensive process with a high cognitive burden. This process would not easily scale with the expected increasing data volume. This competition seeks solutions to challenges in information processing and sensemaking as future ISR moves towards a problem based, deductive and anticipatory approach.

The challenge is not specifically focussed on 'big data' but shares some of the fundamental themes associated with it, eg volume, variety and speed of data. Traditional well-modelled and structured data sources will be augmented by less-structured non-traditional data sources, with different error variables and scales. Different types of information eg military reporting, database of HUMINT, spatial entity locations etc have different variables, because they are formed from distinct data types. They will also have different levels of error and uncertainty as they are collected in unique ways from different sources. An example of an unstructured data source is a text document and a database is a structured data source. The sensitivity and variation of the combined errors of different data types and sources is also important.

MOD is looking to address these challenges by setting up a single intelligence environment to analyse a wide variety of data sources. This will present additional information and intelligence considerations:

- is the information useful in responding to your task?
- is it truthful and accurate, where did it come from?
- does the combination of sources of information support the development of your hypotheses?

The technical problems to be addressed by this themed competition are:

Data association and correlation of both unstructured and structured data

- The association of different types and formats of data sources, with different errors, in different analytical spaces (eg geographic, temporal,semantic) to resolve and aggregate elements.

Uncertainty propagation and management across multiple data representations

- An understanding of error propagation through a variety of data types (reports, photographs, documents), their associated uncertainties and how the data types might be combined ie to direct collection in order to inform or resolve key uncertainties.

Automated hypothesis generation

- Learning lessons from human-centric processes to automatically provide candidate hypotheses, which can then be directly tested against available data sources improving the monitoring modelling and management of intelligence (sensemaking).

Automated learning to understand complex relationships

- The application of machine-learning techniques, against diverse data sets to discover, triage and highlight patterns of interest to the analyst.

Autonomous model generation

- The development of techniques to create machine-generated hypotheses and workflows from the data collected. The models should challenge and provide alternative representations and interpretations for the analysts.

Techniques that cope with large-scale data

- The development of techniques and the application of existing techniques to deal with the challenges in 'big data', paying particular attention to the propagation of error, data accuracy and scalability.

Here are some examples of areas that might produce innovative proposals:

- explore the use of simple financial sector algorithms to the intelligence domain, to improve alerting and notification
- explore options for automated hypothesis generation
- apply previous experience and expertise from commercial sensemaking opportunities to intelligence analysis, such as lessons learnt from Netflix and Nokia Challenges
- create customisable automatic activity dashboards based on the visualisation and analysis of regularly collected datasets
- apply machine-learning techniques to already collected data eg deep learning techniques like Google's word2vec to identify patterns and behaviour within data
- use human and computer co-working on the representation of uncertainty in conceptual models. The model may highlight gaps in collection and so suggest potential future collection requirements
- create user-friendly front ends to models of uncertainty, such as probabilistic graphical models (PGM) that explore the collaborative nature of hypothesis generation and validation ie hypotheses may be developed using a variety of sources which need to be compatible
- examine the scalability and the opportunities within PGM for creation of variables, their dependencies and their values through automatic data mining, machine learning or crowd sourcing
- examine the potential to use advanced information-level fusion algorithms to associate different data sources across different analytical domains eg use the most appropriate spatial analytics as applied to geo-semantic elements within a fused analytical environment.

What we are looking for?

This CDE themed competition is looking for innovative approaches to addressing these challenges. They must scale with the increased data volumes and be more automated in their approach.

Of particular interest are examples of mature academic approaches that have yet to be transitioned to analytical user desktops. We are interested in fast-tracking ideas and academic work to the user community.

The use of open source and open standards will make exploitation of your solution easier. Deliverables must include executable source code, ideally as components in open-source frameworks eg Ozone Widget Framework, so they can be evaluated rapidly. We plan to share feedback of the work with you regularly to aid the in-year development and the future application of the work.

Dstl aims to provide successful bidders with access to a large-scale data set from the VAST 2014 Challenge for testing solutions. Additionally, we will provide a licence to an integrated analytical toolkit for network and geospatial analysis.

Proposals are sought for projects of technology readiness level 3 – 5 (for a description of technology readiness levels see the Acquisition Operating Framework https://www.gov.uk/acquisition-operating-framework) for up to

12 months in duration. We expect solutions to be developed to a maturity where they can take part in the CCS information processing and sensemaking experiment in September 2015 and through on-going exploitation initiatives.

Projects where the principal output is a literature review will not be successful.

The final deliverable should be a follow on proposal for potential Phase 2 funded work. Successful projects will be invited to a stakeholder demonstration event planned for autumn 2015 to support wider exploitation through Dstl and MOD.

Proposers should note that under CDE contracting, they retain ownership of IP and are at liberty to exploit it via other markets.

If deliverables contain background IP this should be clearly identified and you must provide both a full rights and limited version of this deliverable under Defence Contract (DEFCON) 705.

Exploitation

This CDE competition has MOD phase 1 funding from the following projects:

- C4ISR Concepts and Solutions
  - To inform future MOD C4ISR capability planning the MOD wants to develop and validate credible C4ISR concepts at the enterprise level, looking across the DCPD Intelligence cycle and performing assessment and demonstration of whole system solutions. This work is being performed under the C4ISR Concepts and Solutions (CCS) project (formerly ISTAR Concepts and Solutions, ICS) developing and validating credible options for future C4SIR systems, looking out to the next 3 to 30 years.

- Intelligence Collection and Exploitation
  - The Intelligence Collection and Exploitation (ICE) project is a project within the Command Control Information and Intelligence (C2I2) S&T programme. The ICE project innovates, explores, develops and transitions technology solutions across the multi-INT space applied to real data and real problems for a variety of key stakeholders. The ICE project works across both applied and core R&D angles focussing on text analytics, spatio-temporal correlation and data association. The applied research strand focuses on current military operations, work in a data rich environment and to not lose sight of the art of the possible. ICE uses open source technology where possible and transitions technology quickly to operations (whilst considering the full range of system issues necessary to realise a military capability).

We expect to take this competition topic forward as an enduring requirement within both projects; it is our aim the most successful projects will be taken forward for phase 2 funding. Up to £1M will be available for phase 2 funding.

MOD exploitation of successful phase 1 work has two possible routes:

- through the CCS Information Processing and Sensemaking project
- transition to beta-style[2] "silver level" environments through a *developing MOD technology transition model*[3] as part of ICE

[2] https://www.gov.uk/service-manual/phases/beta.html

[3] The model enables the development of code and web applications at lower levels. Potentially useful applications can have real users, real data and real problems worked against them, creating qualitative and quantitative user feedback to support that technology transition. It ensures tools are relevant and tested before deployment to Military Operations, enabling the management of risk.

Invitation for CDE proposals

This competition will be supported by presentations given at the launch seminar on 22 May 2014. These will be available to download via the competition page

Proposals are invited from industry and academia in the UK and overseas for research that can demonstrate a proof-of-concept to meet one or more of the challenges for "**Information Processing and Sensemaking**".

The total available for funding for this competition is £600k.

There is no cap on the value of proposals but it is more likely that a larger number of lower value proposals (eg £50k - £100k) will be funded than a small number of high value proposals.

Proposals should focus on delivering early and often throughout the contract which will last up to 12 months in duration. Main 'source code' deliverables must be completed by August 2015, for a demonstration in September 2015 potentially using data provided by Dstl. Proposals should include a descriptive scoping for a longer programme of follow on work of any duration. However the proposal must be clearly partitioned with a costed proof-of-concept stage which is the focus of this CDE competition. Proposals for work beyond the proof-of-concept stage will only be considered after the proof-of-concept stage has delivered, using the understanding gained to make an informed decision.

Proposals must include:

- a clear statement of what challenge the solution is aimed towards
- a clear description of what is novel and innovative in the solution
- a clear statement of the programme of work that would be carried out and the outputs (deliverables) from the work
- a clear statement of the expected outcome(s), how this will be proven or demonstrated and how it will provide evidence that the outputs can be exploited
- a clear description of the value of the solution to operational capability including the likely saving to through-life costs
- a statement on the anticipated practicality of adopting the proposed solution
- an outline of any data/equipment requirements of the proposal, and how these will be met. Any dependencies on the supply of data/equipment from MOD must be stated
- confirmation that the source code will be delivered.

Proposals that do not include the required information will not be successful.

Proposals will be assessed by subject matter experts from MOD and Dstl using the MOD Performance Assessment Framework (PAF). Deliverables from contracts will be made available to Technical Partners and subject to review by UK MOD.

Dstl will be available to provide advice and/or guidance via an appointed Technical Partner throughout the project and provide the interface with MOD and wider Government stakeholder community.

Dstl does not commit to fund any follow on work as a result of any contracts placed via this CDE competition, but more promising ideas will be considered for further funding where appropriate.

CDE proposal submission process

Key dates

- 22 May 2014 Competition launch event at Altitude 360, Milbank Tower, London
- 3 June 2014 Post-launch webinar
- 26 June 2014 Competition closes at 1700 hrs
- 31 August 2015 Proof-of-concept research complete
- September- November 2015 1 week demonstration at CCS trial

We expect that contract placement will be started in early August 2014 and feedback to unsuccessful bidders issued in mid-August 2014.

Proposals for funding must be submitted by 1700 hrs on Thursday 26 June 2014 using the CDE portal. Proposals must be clearly marked with "Information Processing and Sensemaking" as a prefix in the title.

Please plan the timeline for submitting your proposal carefully. If you have not used the CDE portal before you will need to read through the guidance, including how to open an account starting with the Quick Start Guide.

Other information and guides are available on the CDE website:

- general CDE advice: www.science.mod.uk/engagement/cde/working_with_cde.aspx
- contract & IPR guidance: www.science.mod.uk/engagement/cde/funding_contracts.aspx
- on using the portal: www.science.mod.uk/engagement/the_portal.aspx. The portal is optimised for proposals based on physical sciences and engineering and we are aware that proposers sometimes struggle to adapt to using it with social science based proposals. The key points (rather than the detailed questions) that are sought under the main headings still apply and further advice can be obtained from CDE.

Common errors in preparing and submitting a proposal include:

Character limit – there is a limit of 1000 characters in each individual descriptive paragraph within the proposal; when completed they must be added to the document; additional paragraphs can be added if 1000 characters is insufficient.
It is a web-based tool – please save your work regularly to avoid 'time-outs' that lose work**attachments fail** – they must be Word 97-2003 , portrait format, have generous margins with no material overhanging the margin and a max size of 1 MB. Please note that attachments should only be used for supplementary information, the main points of your proposal should be written into the online form. Care should also be taken to make sure that attachments are placed in the relevant section (eg technical information should not be attached to the commercial section).
Failing to properly submit - **publish is not the same as submit**. You have **not** completed the submission process if your proposal is at the FINAL/PUBLISHED stage (in the status and published status columns respectively); CDE has no sight of the proposal at this stage. To complete submission you need to press 'submit' under the 'Tasks' column. This changes the status of your proposal to 'SUBMITTED'; it will then change (normally after a few days, often sooner) to 'RECEIVED' indicating that the proposal has been accepted by CDE for assessment.

For a proposal to be accepted for assessment:

- the standard terms and conditions of CDE must be unequivocally accepted
- there must be at least one deliverable against which payment can be made
- the commercial section of the proposal must be completed.

Please do not leave submission of your proposal until just before the deadline. Past experience has shown that the portal becomes heavily loaded near a competition close resulting in slow operation (up to one hour to publish rather than a few minutes) and that, with the pressure of the deadline, mistakes are made that mean proposals are not submitted or accepted.

Proposals and content placed on the portal should not contain classified information.

Queries and help

As part of the proposal preparation process, queries and clarifications are welcomed:

Technical queries about this specific competition should be sent to **dstlsensors@dstl.gov.uk**.
Capacity to answer these queries is limited in terms of volume and scope. Queries should be limited to a few simple questions or if provided with a short (few paragraphs) description of your proposal, the technical team will provide, *without commitment or prejudice*, broad yes/no answers. This query facility is not to be used for extensive technical discussions, detailed review of proposals or supporting the iterative development of ideas. Whilst all reasonable efforts will be made to answer queries, CDE and Dstl reserve the right to impose management controls when higher than average volumes of queries or resource demands restrict fair access to all potential proposal submitters.

General queries (including how to use the portal) should be sent directly to CDE at cde@dstl.gov.uk.