SECTION 1: OPPORTUNITY DESCRIPTION

1.A Program Overview

The Rapid Explanation, Analysis, and Sourcing Online (REASON) Program aims to develop technology that will enable intelligence analysts to substantially increase the quality of argumentation in their analytic reports through more effective use of evidence and reasoning. In the context of an analytic report, evidence is information that supports a judgment, while reasoning is the stated justification for the judgment. The technology developed by the REASON Program will automatically produce comments (feedback and recommendations) on a draft report, highlighting additional relevant evidence, and identify strengths and weaknesses in the draft's reasoning. Analysts can use the comments to improve their reports.

Argumentation is central to the Intelligence Community (IC) Analytic Standards, which are listed in Intelligence Community Directive (ICD) 203[1]. The standards are intended to guide IC analysis and analytic production. The Analytic Tradecraft Standards focus on several aspects of evidence and reasoning, including sourcing, explaining uncertainty, distinguishing between underlying information and assumptions, and logical argumentation. Because evidence and reasoning are crucial components of every analytic report, REASON will have broader application than previous research efforts aimed at helping the IC make accurate forecasts.

Currently, intelligence analysts are encouraged to use structured analytic techniques to boost the quality of argumentation in their reports.[2,3] Many of these methods require substantial additional quantities of analysts' time and are therefore not widely used. As contrasted with current applications of structured analytic techniques, REASON technology will automatically produce comments with no additional effort from analysts, who can use any comments they find valuable. Some of these comments might be based on the automated application of effective structured analytic techniques, along with additional innovations.

By making specific suggestions on draft analytic reports, REASON technology will fit into the existing intelligence analysts' workflow. The suggestions will be analogous to those made by automated spelling and grammar checks, except that REASON's suggestions will focus on improving argumentation instead of writing.

Offerors shall address three technical Task Areas (TAs) to meet REASON's goal of developing automated methods to produce comments on draft analytic reports that enable analysts to substantially increase the report's quality of argumentation:

- **Task Area 1 (TA1) – Identify Additional Evidence:** Automatically find relevant supporting and contrary evidence in addition to the evidence used in a draft report.
- **Task Area 2 (TA2) – Identify Reasoning Strengths and Weaknesses:** Automatically find strengths and weaknesses in the reasoning of a draft report.
- **Task Area 3 (TA3) – Produce Recommendations to Increase Quality of Argumentation:** Based in part on the output of TA1 and TA2, automatically produce comments that enable analysts to substantially improve the argumentation in their reports.

Offerors must propose novel approaches to each of these three TAs, and if selected as a Performer, will be required to create an end-to-end technology that incorporates software components from each TA. Developed capabilities must be compatible with a provided Application Programming Interface (API) to facilitate assessment by independent test and evaluation (T&E) according to program metrics described in Section 1.F, Program Metrics.

---

[1] https://www.dni.gov/files/documents/ICD/ICD%20203%20Analytic%20Standards.pdf.
[2] https://psu.pb.unizin.org/app/uploads/sites/65/2018/01/Tradecraft-Primer-apr09.pdf
[3] https://www.dia.mil/FOIA/FOIA-Electronic-Reading-Room/FileId/161442/

1.A.1 Technical Challenges and Objectives

Offerors shall address the following technical challenges and objectives to meet the REASON goals.

Identify Additional Evidence (TA1): The goal of TA1 is to develop technology that automatically identifies additional supporting and contrary evidence when such evidence exists. Successful approaches will produce, in response to a draft analytic report and a corpus of source documents, a prioritized list of up to eight[4] items of additional evidence contained in the corpus but not mentioned in the draft report. (Only the first eight items will be scored). Performer systems will need to determine whether a piece of information is relevant evidence bearing on the analytic question addressed in the draft report and whether it is additional (non-redundant) to the information used in the draft report. Performer systems should identify (where appropriate) contrary as well as supporting evidence, determined by the relationship of the evidence to either the draft's conclusion or the evidence and reasoning within the draft. When there is no non-redundant additional evidence, performer systems should report that.

Identify Reasoning Strengths and Weaknesses (TA2): The goal of TA2 is to develop technology that automatically finds strengths and weaknesses in the reasoning of a draft analytic report. This means that the system will be able to identify reasoning elements in the draft report. For each draft report, the system will identify up to eight strengths and weaknesses. (Only the first eight items will be scored). Each strength or weakness will point to the appropriate section of the draft report and shall be accompanied by a brief explanation of why it is strong or weak reasoning. A successful system must distinguish between apparent and real strengths and weaknesses in reasoning. If the reasoning in the draft analytic report is sound, the system will report that.

Produce Recommendations to Increase Quality of Argumentation (TA3): The goal of TA3 is to develop a software application that, with input from TA1 and TA2, automatically produces comments on draft analytic reports that enable analysts to substantially increase the quality of argumentation in the report. TA1 and TA2 provide evidence and reasoning improvement as inputs to TA3, and TA3 builds on these to present useful suggestions to the analyst. Successful approaches must identify key areas where a draft report can be improved and generate comments based on those, presented in a manner that prompts analysts to use them effectively. The comments may concern individual issues or the overall draft report, including the correctness of the conclusion or the appropriateness of cited evidence; they may address the content or the communication of the report's argument.

The TA3 REASON application deliverable must be compatible with analyst authoring applications, typically Microsoft Word. TA3 encompasses both the software development and the research necessary to draw on TA1 and TA2 inputs and effectively communicate recommendations to the analyst so that they improve the argumentation in their draft report.

1.A.2 Program Phases

The REASON Program is a 42-month effort, comprising two phases. Proposers must submit to both phases or else they will be considered to be non-compliant. Because the goal of the REASON Program is to increase the quality of reports produced on classified systems, deliverables produced by proposers must grant the Government intellectual property (IP) rights sufficient to allow the Government to modify and deploy deliverables on classified networks.

In Phase 1 performer systems will be tested comprehensively on unclassified data consisting of draft analytic reports and news reports. In parallel, the REASON independent T&E team will retrain and evaluate performer systems on classified draft analytic reports and source reports. In Phase 2 performer systems will be tested on classified data consisting of draft analytic reports and classified source reports. In each phase performers will develop and be tested on techniques for addressing TA1, TA2, and TA3. Each phase

[4] The limit of eight was chosen based on balancing comprehensiveness, avoidance of overwhelming analysts with feedback, and practical testing considerations.

will contain several testing cycles; each cycle will contain approximately 20 challenge problems consisting of an analytic question and a draft report. Challenge problems will become increasingly difficult over the course of a phase.

In each phase TA1 will be evaluated by measuring the performer system's ability to automatically find and rank-order additional evidence. TA2 will be evaluated by measuring the performer system's ability to automatically find and explain strengths and weaknesses in reasoning. TA3 will be evaluated in two ways:

1. T&E raters will evaluate the comments produced by performer systems – on correctness, appropriateness, and clarity.
2. Final Exam: Human participants will be assigned to use a performer system or to be in a control group. Participants will produce draft reports on assigned analytic questions and have opportunities to revise the reports. Participants assigned to a performer system will see the comments that the system produces and may use any of the comments in revising their report. T&E will measure the argumentative quality of finished reports, comparing those produced with the aid of a performer system to those produced by participants in the control group.

1.A.2.1 Phase 1

Phase 1 shall have a duration of 24 months. The goal of Phase 1 is to develop novel systems to enable analysts working with unclassified data to produce analytic reports of substantially higher quality. Performer research will focus on developing automated methods for processing argumentation (evidence and reasoning) accurately, producing comments that human users find explainable and helpful.

Performer systems' TA1 and TA2 capabilities will be tested over three cycles each, where each cycle includes both unclassified and classified testing. Cycles will become increasingly difficult during the phase: it will become more challenging to find additional evidence and strengths and weaknesses in the reasoning. Performer systems' TA3 capabilities will be tested in one cycle (with unclassified and classified testing) and one final exam using unclassified data. Each of the TA3 cycles will measure systems' ability to produce comments that are correct, appropriate, and clear. The final exam will measure the effect of the system's automatically produced comments on the quality of reports written by human users who produced the draft and can view the comments. The human users will include undergraduate or graduate students in disciplines such as intelligence analysis or international relations.

All Performer work will be unclassified. Performer systems will be tested by cleared T&E personnel on classified data but Performers will not be able to review that classified data. T&E will provide Performers with unclassified summary results from classified testing. In classified testing, Performer systems, operated by cleared T&E personnel in an automated fashion, will need to search, identify, and process textual documents containing classified data. These documents will differ from the unclassified news and opinion documents in several ways. In addition to containing new information, the classified data will have distinctive stylistic features, including classification markings and IC-specific jargon and abbreviations. Some unclassified examples with these stylistic features will be provided at Program Kickoff.

1.A.2.2 Phase 2

Phase 2 shall have a duration of 18 months. The goal of Phase 2 is to refine the capabilities of the methods developed in Phase 1 so that they function effectively on classified data and produce substantially larger effects. Performers will refine their systems to process the content and style of the IC's source reports using unclassified examples, but they will not have access to classified data. Performers will receive actionable summary level unclassified feedback from the independent cleared T&E that they can use to refine the capabilities of their methods and systems.

Performer systems' TA1 and TA2 capabilities will be tested over one cycle each. Performer systems' TA3 capabilities will be tested over two cycles and one final exam, using cleared intelligence analysts as participants.

1.B Team Expertise

Collaborative efforts and teaming among Offerors are highly encouraged. It is anticipated that teams will be multidisciplinary and may include expertise in one or more of the disciplines listed below. This list is included only to provide guidance for Offerors; satisfying all the areas of technical expertise below is not a requirement for selection, and unconventional or innovative team expertise may be needed based on the proposed research. Proposals should include a description and the mix of skills and staffing that the Offeror determines will be necessary to carry out the proposed research and achieve Program metrics.

- Applied epistemology
- Argumentation
- Cognitive psychology
- Experimental design
- Informal logic
- Judgment and decision making
- Linguistics
- Natural language processing
- Philosophy of language
- Psychometrics
- Rationality
- Software engineering
- Systems engineering
- Systems integration

1.C Program Scope and Limitations

Proposals shall explicitly address all the following:

- **Underlying Theory:** Proposed strategies to meet Program-specified metrics must have firm theoretical bases that are described with enough detail that reviewers will be able to assess the viability of the approaches. Proposals shall properly describe and reference previous work upon which their approach is founded.
- **R&D Approach:** Proposals shall describe the technical approach to meeting Program metrics.
- **Technical Risks:** Proposals shall identify technical risks and proposed mitigation strategies for each.
- **Software Development:** Proposals shall describe the approach to software architecture and integration.

The following areas of research are **out of scope** for the REASON Program:

- Purely automated production of analytic reports.
- Approaches that process non-textual inputs such as:
    - Images
    - Video
    - Audio
    - Structured data sources

    - However, it is permissible to use textual clues (e.g., image captions) to locate and retrieve non-textual items
- Approaches aimed at processing text in languages other than English.
- Approaches that require Performer access to classified information or data. **All** Performer research will be strictly unclassified.

1.D Program Data

The REASON program will use both data provided by the Government Team and data provided by Performers. Proposals must specify the data needed to carry out the proposed research and what data characteristics are necessary for the Proposer's approach(es) to be successful at meeting program objectives. These details should be provided for using Government-provided data as well as Performer-provided data.

1.D.1 Government-Provided Data

The Government will obtain data as a corpus of source documents and make it available to Performers via a T&E testbed. At the beginning of Phase 1 this will be a corpus of unclassified news articles and analytic reports. At the beginning of each phase the Government team will also provide access to a small sample of unclassified draft reports similar in form to the draft reports that will be used in T&E testing in that phase.

1.D.2 Performer-Provided Data

Each performer is expected to have a unique technical solution to the REASON challenges and may require additional data for model training, model running, internal evaluation, or other research needs. Proposals must present a dataset development plan detailing how the team intends to obtain the data required. This documentation should account for any other associated labor to curate and facilitate use of data that are acquired.

As part of their proposal, each team shall prepare a REASON Privacy Plan Version 1.0 that comprehensively describes the efforts the team will take to protect personally identifiable information and safeguard the security of any personal data collected or services involved in collection, transmission, processing, and storage of these data. Any claims that data are anonymous must be based on evidence and supported with sufficient information regarding how the data have been anonymized.

This version 1.0 of the REASON Privacy Plan shall be included in the Proposer's proposal as Attachment 6 that covers all external datasets to be leveraged as part of the proposed research approaches. The REASON Privacy Plan shall be updated at the beginning of each Phase and when new sources of data or datasets are proposed for use within a Performer's REASON research activities, including data used for either development or evaluation purposes.

1.E Test and Evaluation (T&E)

T&E will be conducted by an independent team of contractor staff carrying out evaluation and analyses of Performer research deliverables using program test datasets and protocols. In addition to independent T&E, the program will regularly gauge interim progress of Performer research activities towards REASON objectives and target metrics using T&E results measured and reported by the Performer teams themselves.

The REASON Program will pursue rigorous and comprehensive T&E to ensure that research outcomes are well characterized, deliverables are aligned with program objectives, and performance is measured across the full range of conditions. T&E activities will inform IARPA and Government stakeholders on REASON research progress and serve as invaluable feedback to Performers to improve their research approaches, training practices, and system development.

Performers will have specific Deliverable Milestones driven by the REASON evaluation cycle schedule at which all subcomponent and system algorithms and software will be delivered to IARPA and its designated T&E Team. The T&E Team will then conduct independent evaluations with the objective of characterizing the quality, functionality, and performance of the REASON systems. In addition to quantitative measurements, T&E assessments will be carried out to establish a thorough understanding of the progress, status, and limitations of the Performer's research.

For classified testing, the T&E Team will retrain Performer systems to classified data as necessary, using scripts or processes provided by the Performers.

T&E results and feedback will be provided to Performers at regular intervals to keep them abreast of current independent performance measurements and to inform and improve their R&D approaches and methods. T&E will provide unclassified feedback summarizing the results of the unclassified testing and the classified testing to Performers. T&E results from all Performers will be shared with all teams to establish an understanding of the current state and progress of REASON research; T&E results will also be shared with USG external stakeholders, including their contractors, for Government purposes. IARPA may conduct other supplemental evaluations or measurements at its sole discretion to evaluate the Performers' research and Deliverables.

1.F Program Metrics

Achievement of metrics is a performance indicator under IARPA research contracts. IARPA has defined REASON program metrics to evaluate effectiveness of the proposed solutions in achieving the stated program goal and objectives, and to determine whether satisfactory progress is being made. The metrics described in this BAA are shared with the intent to scope the effort, while affording maximum flexibility, creativity, and innovation to Proposers proposing solutions to the stated problem.

The REASON T&E protocols and evaluation methodology are currently under development; further details will be provided at Program Kickoff in the Phase 1 REASON T&E Plan. Program metrics may be refined during the various phases of the REASON program; if metrics change, revised metrics will be communicated in a timely manner to Performers. The evaluation methodology may be revised by the Government at any time during the program lifecycle to better meet program needs. The preliminary program metrics and target scores are provided below.

The TA1 metric is a modified version of alpha normalized discounted cumulative gain (α-nDCG), which will use the union of the outputs (evidence items) from all Performers and combine it with the outputs of a manual search for evidence by the T&E team in order to approximate the ideal results. The formula for α-DCG will be:

$$\text{anDCG} = \frac{\text{aDCG}(\text{Performer results})}{\max(\text{aDCG}(\text{All Performer results} \cup \text{T\&E discovered items}))}$$

The scoring process will be:

1. T&E performs a manual search of the corpus for evidence at the time they create each TA1 Challenge Problem. Some of their search results will be cited in the draft report, which is used as the input to Performer systems; others will be reserved but not cited.
2. The Performer systems produce a set of up to 8 ordered evidence items found in the corpus as the output for the Challenge Problem.
3. The output evidence item result sets from the several Performer systems are combined with the reserved evidence items from step 1. The same item may be returned by multiple Performers or may match the T&E items.
4. Each of the items in the resulting set from step 3 are evaluated:

a. Is the item redundant to the cited evidence in the draft report?
b. What is the relevance (score of 1, 2, or 3) of the item?
c. What category does the item belong to? For example, if the analysis report deals with a potential military invasion, then one category might include evidence of troop movements, another might include public statements by leaders, a third might include previous examples of similar circumstances, etc. The categories will be used to calculate the diversity of the cited evidence.

5. For each candidate subset of 8 items from the result set compiled in step 3 as assessed in step 4, determine the α-DCG. Take the maximum value as the denominator for computing α-nDCG.

TA2 has two metrics: Reasoning Explanation Quality (REQ) and F1. REQ will assess the explainability of the identified strengths and weaknesses. T&E raters will evaluate the correctness and clarity of each explanation of a reasoning strength or weakness on a 1-4 scale.

T&E will measure Performer system's identification of strengths and weaknesses in reasoning of draft report evaluated using F1 Score, which gives credit for two features:

- If the system says X is a strength or weakness, is the system correct (*i.e.* is the system output a true positive?) or is it wrong (*i.e.* the system output is a false positive).
- If X is a strength or weakness, does the system says so? (If not, then the system output is a false negative)

TA3 has two metrics. The first is REASON Comment Quality (RCQ). T&E raters will evaluate the comments provided by the Performer TA3 system. RCQ scores will be based on correctness, appropriateness, and clarity of the comments, using a 1 (poor) - 4 (excellent) scale.

The second TA3 metric applies to the final exam. The finished analytic reports produced by the human participants will be evaluated by T&E raters using Report Quality Score (RQS). RQS is based on scores of six of the IC Analytic Tradecraft Standards: sourcing, uncertainty, assumptions, alternatives, logic, and accuracy. Each finished report will be graded on each standard, with a range from 1 (poor) to 4 (excellent), so RQS values range from 6 to 24. Each performer system's RQS will be compared to the control group's RQS.

A summary of metric targets by Phase is shown in Table 1; these are subject to change over the course of the program. Final Phase 1 metrics will be presented at kickoff.

Table 1: REASON Program Target Metrics

Task	Metric	Phase 1 Target	Phase 2 Target
TA1: Identify Additional Evidence	**α**-nDCG	> 0.25	> 0.40
TA2: Identify Reasoning Strengths and Weaknesses	Reasoning Explanation Quality (REQ)	> 2.75	> 3.5
	F1	> 0.65	> 0.80
TA3: Produce Comments to Increase Quality of Argumentation	REASON Comment Quality (RCQ)	> 2.75	> 3.5
	Report Quality Score (RQS)	ΔRQS > 1.5	ΔRQS > 3.0

1.G Program Waypoints, Milestones, and Deliverables

Waypoints, Milestones, and Deliverables are established from the Program's onset to ensure alignment with REASON objectives, organize research activities in a logical and reportable manner, and facilitate consistent and efficient communication among all stakeholders – IARPA, REASON T&E, USG Stakeholders, and Research Performers. A schedule of key program milestones and deliverables in shown in Figure 1.

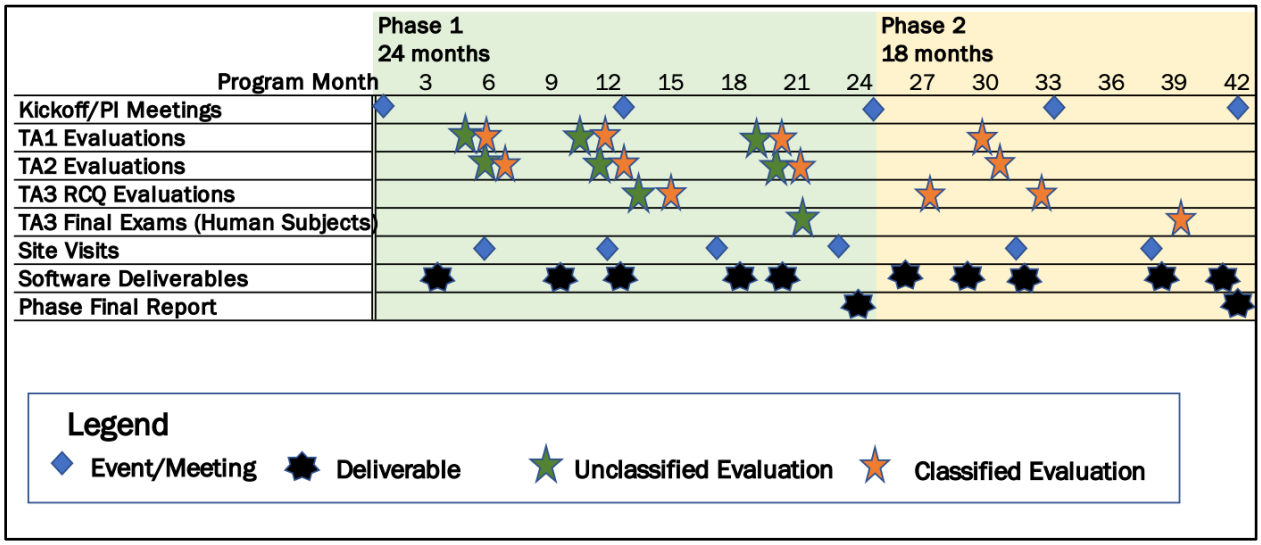

Figure 1. Schedule of Key Milestones and Deliverables

1.G.1 Program Milestone, Waypoint, and Deliverables Timeline

Phase	Month	Event	Description	Comment	Deliverable
1-2	All	Waypoint	Monthly Status Report	Due on 15th of each month	MSR
1-2	All	Waypoint	Progress and Status Meeting	Monthly teleconference with REASON PM	N/A
1	1	Waypoint	Kickoff Meeting	DC Metro Area	Presentation Materials
1	1	Waypoint	Sample Data	Provided as GFI	N/A
1	4	Waypoint	Site Visit	Performer Site	N/A
1	7	Deliverable	TA 1 and 2, Cycle 1	Performer system output and software	Software Container
1	12	Deliverable	TA 1 and 2, Cycle 2	Performer system output and software	Software Container
1	10	Waypoint	Site Visit	Performer Site	N/A
1	15	Deliverable	TA3, Cycle 1	Performer system output and software	Software Container
1	13	Waypoint	PI Meeting	DC Metro Area	Presentation Materials
1	16	Waypoint	Site Visit	Performer Site	N/A

Phase	Month	Event	Description	Comment	Deliverable
1	18	Deliverable	TA 1 and 2, Cycle 3	Performer system output and software	Software Container
1	19	Deliverable	TA3, Final Exam	Performer system output and software	Software Container
1	22	Waypoint	Site Visit	Performer Site	N/A
1	24	Deliverable	Phase 1 Final Software Delivery	Performer system output and final Phase 1 software	Software Container
1	24	Deliverable	Phase 1 Final Report		Report
2	25	Waypoint	Kickoff Meeting	DC Metro Area	N/A
2	25	Waypoint	Sample Data	Provided as GFI	N/A
2	27	Deliverable	TA3, Cycle 2	Performer system output and software	Software Container
2	28	Waypoint	Site Visit	Performer Site	N/A
2	29	Deliverable	TA 3, Cycle 4	Performer system output and software	Software Container
2	31	Deliverable	TA 1 and 2, Cycle 4	Performer system output and software	Software Container
2	32	Waypoint	PI Meeting	DC Metro Area	Presentation Materials
2	33	Waypoint	Site Visit	Performer Site	N/A
2	34	Deliverable	TA 3, Cycle 3	Performer system output and software	Software Container
2	38	Deliverable	TA 3, Final Exam	Performer system output and software	Software Container
2	39	Waypoint	Site Visit	Performer Site	N/A
2	41	Waypoint	PI Meeting	DC Metro Area	Presentation Materials
2	42	Deliverable	Phase 2 Final Software Delivery		Software Container
2	42	Deliverable	Phase 2 Final Report		Report

1.G.2 Software Deliverable Formatting

Performers will be required to provide algorithm and software deliverables (including source code and executables) in a manner that conforms to a standardized industrial method or methods that will be provided at Program Kickoff. To facilitate planning, Offerors may assume that the standardized configuration will require the use of software containerization technology (e.g., Docker and a REST API). This means that the entirety of a Performer's system, including pre- and post-processing, must be included within the delivered software container. These systems must be able to accept inputs in the form of Challenge Problems from an API to be developed by the T&E team and to submit outputs to that API.

For software that includes models that require initial training, the expectation is for the initial model training to occur on Performer systems, with the ability for the T&E Team to re-train and test the model with the same and/or other data.

If Offerors plan to use cloud computing resources for model development and training, they should include descriptions of these requirements in their technical approach descriptions. Retraining of Performer systems for T&E purposes will be subject to limitations on system retraining time and resources. Those limitations will be briefed at Program Kickoff.

Each team is required to include among their Key Personnel a Lead System Integrator (LSI) who shall be responsible for preparing software Deliverable subcomponents, modules, and systems, performing quality control of Deliverable, and integrating key components into the primary REASON system(s). The LSI will also oversee communication and coordination across a Performer's research teams including subcontractors, if applicable, to ensure that research products are functional, integrated and following software coding best practices (e.g., inline comments, documentation). Additional team members and roles are dependent on the proposed research, as such, there is no predetermined or required skill mix.

1.G.2.1 Program API

The REASON Program will use a standardized API for all software deliverables and evaluations. The first version of the REASON API will be provided to Performers at the Phase 1 Kickoff Meeting and updated periodically thereafter. The API will define function calls, data structures, and gallery creation and management for operating and evaluating REASON software in a standardized manner. The API will be functionally identical for unclassified and classified testing. Specifically, the API will provide access to the document corpus for automated retraining of Performer systems, delivery of the Challenge Problems used in T&E evaluations, and submission of result sets for Challenge Problems.

1.G.3 End of Phase Final Reports

At the end of each Program Phase Performers will be required to submit a comprehensive Final Report that describes their efforts and results during the Phase. These reports shall include an executive summary, a description of the technical approach taken, details on the results, findings, and technical insights gained from the R&D effort, lessons learned, and suggested future research directions. The Final Report shall also include high level system design documentation for the final software deliverable. This design documentation shall include any hardware requirements and dependencies on third-party software libraries.

1.H Meeting and Travel Requirements

Performers are expected to assume responsibility for administration of their projects and to comply with contractual and program requirements for reporting, attendance at program workshops, and availability for site visits. The following paragraphs describe typical expectations for meetings and travel for IARPA programs as well as the contemplated frequency and locations of such meetings. In addition to ensuring that all necessary details of developed software, algorithm, and operational instructions are clear and complete, each Performer will be required to be available for questions and troubleshooting from the T&E Team in weekly and/or bi-weekly status meetings.

1.H.1 Workshops

All Performer teams are expected to attend workshops, to include Key Personnel from prime and subcontractor organizations.

The REASON Program intends to hold a program Kickoff Meeting workshop in the first month of the program and first month of the subsequent program phase. In addition, the program will hold a PI Review Meeting at the end of each phase and at the phase midpoint. Kickoff Meetings and PI Review Meetings may be combined for logistical convenience.

Both types of meetings will likely be held in the Washington, D.C. metropolitan area, but IARPA may opt to co-locate the meeting with a relevant external conference or workshop to increase synergy with stakeholders. IARPA reserves the right to hold the meeting virtually for logistical or health and safety reasons.

Kickoff Meetings will typically be one day in duration and will focus on plans for the coming Phase, Performer planned research, and internal program discussions. PI Review Meetings will typically be two days in duration and will have a greater focus on communicating program progress and plans to USG stakeholders. These meetings will include additional time allocated to presentation and discussion of research accomplishments.

In both cases, the workshops will focus on technical aspects of the program and on facilitating open technical exchanges, interaction, and sharing among the various program participants. Program participants will be expected to present the technical status and progress of their projects to other participants and invited guests. Individual sessions for each Performer team with the REASON Program Manager and the T&E Team may be scheduled to coincide with these workshops. Non-proprietary information will be shared by Performers in the open meeting sessions; proprietary information sharing shall occur during individual breakout sessions with the REASON Program Manager and the T&E team.

1.H.2 Site Visits

Site visits by the Government Team will generally take place semiannually during the life of the program. These visits will occur at the Performer's facility and last no longer than two days. Reports on technical progress, details of successes and issues, contributions to the program goals, and technology demonstrations will be expected at such site visits. IARPA reserves the right to conduct additional site visits on an as-needed basis.

1.I Period of Performance

The REASON program is envisioned as a 42-month effort that is intended to begin October 1, 2023.

Phase 1 (Base Period): October 1, 2023 – September 30, 2025

Phase 2 (Option 1): October 1, 2025 - March 31, 2027

1.J Place of Performance

Performance will be conducted at the Performers' sites.

1.K Security

All Performer work will be unclassified. Performer systems will be tested using classified data but Performers will not be able to review that classified data. Performers will be provided with unclassified summary results from classified testing.