

#	Question	Answer
1	<p>There are two Section 3s in the BAA.</p> <p>Section 3: Program Phases Section 3: Recommended Team Expertise</p> <p>Should "Recommended Team Expertise" be labeled as Section 4 and subsequent sections renumbered accordingly?</p>	<p>Yes, the government recognizes this administrative error, and applicants should continue to work with the content provided in those sections.</p>
2	<p>For FAR Clause 52.227-14, since we are an institution of higher education can we request that Alt IV be used?</p>	<p>Per Section II.E.2. of the BAA Amendment 1 (pg 53-54), negotiations will only take place with Offerors who are selected for negotiation.</p>
3	<p>For FAR Clause 52.204-14, since we are an institution of higher education we can request this to be removed as it does not apply to us?</p>	<p>Per Section II.E.2. of the BAA Amendment 1 (pg 53-54), negotiations will only take place with Offerors who are selected for negotiation.</p>
4	<p>We understand that federal [Department of Energy (DOE) National Labs] cannot be funded as part of a performer team in an IARPA program such as REASON. However, a member of our team is a small business based in [a state where there is a Small Business Assistance Program that] has an arrangement to make technical resources at [specific] National Labs available to small businesses in the state at no cost to the business, and [these National Labs have] resources that our partner wishes to access as part of the project we expect to propose. We find this arrangement attractive because it reduces the cost of our contribution to IARPA, and also may lead to transition opportunities. However, it leads to two questions. a) Given that our proposal will not include any funding for [the National Labs], are we correct in concluding that the described arrangement is compliant with IARPA's restrictions on FFRDC participation? b) We understand that REASON will utilize a T&E team that may include participants from Federal labs. If [personnel from these Federal Labs] are part of the REASON T&E team, may our small business teammate collaborate with other [Federal Lab] personnel, assuming that a firewall is implemented between our collaborators and the [Federal Lab's] REASON T&E team members?</p>	<p>Per Section II.C.1 of the BAA Amendment 1 (pg 21), Government-Owned, Contractor-Operated facilities like DOE National Labs "have a special relationship with the Government, that gives them access to privileged and/or proprietary information or access to Government equipment or real property [and] are not eligible to submit proposals under this BAA or participate as team members under proposals submitted by eligible entities."</p>

5	<p>In the BAA it states, "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 II.A.7, Program Metrics." We were wondering if participating with a Performer and providing the T&E platform, assuming that will be a separate TO likely competed via another ODNI contracting vehicle, would be mutually exclusive?</p> <p>Appreciate your guidance if it's a one or the other type of situation, or if we can submit proposals for both!</p>	<p>Offerors shall follow the instructions regarding organizational conflict of interest in Section 1.A.1 Organizational Conflict of Interest (OCI) (see page 22-23 of the amended BAA), including notifying the Government regarding any potential OCIs.</p>
6	<p>Are research assistants in a professional faculty research lab considered advisees? Are students considered advisees?</p>	<p>Per the requirement in Volume 1 Part E "Project Contributors" (pg. 27), it is the responsibility of the Offeror to identify all anticipated Project Participants and define the organizational structure, functional roles, and tasks.</p>
7	<p>Are supervisors in a professional faculty research lab considered advisors?</p>	<p>Per the requirement in the BAA Amendment, Volume 1 Part E "Project Contributors" (pg. 27), it is the responsibility of the Offeror to identify all anticipated Project Participants and define the organizational structure, functional roles, and tasks.</p>
8	<p>Can you clarify potential restrictions on involvement of Chinese citizens with Green Cards working in the US, South Korean citizens working in South Korea, British citizens working in South Korea and Israeli citizens with Green Cards working in the US?</p>	<p>Offerors should review the Eligible Applicants section of the amended BAA in Section II.C.1 (pg 21-22).</p>
9	<p>If foreign nationals are members of teams, what specific requirements for ensuring security of methods and data will be expected?</p>	<p>Offerors should review the Eligible Applicants section of the amended BAA in Section II.C.1 (pg 21-22)</p>
10	<p>p. 21 of the BAA says: "Eligible applicants under this BAA include Institutions of higher education (foreign and domestic), nonprofit organizations, and for-profit concerns (large and small businesses)." I have a few questions about this:</p> <ol style="list-style-type: none"> 1. Does this mean that independent researchers who are not affiliated with such an entity are ineligible to be part of an Offeror team? 2. Does it mean that individuals who are affiliated with such an entity--e.g., a university--must have that entity officially be part of the Offeror team? 3. What if an individual is affiliated with such an entity but is not employed by them? How should such an individual present themselves as part of an Offeror team? 	<p>Per Section II.C.1 in the Amended BAA (pg 21), "all responsible sources capable of satisfying the Government's needs may submit a proposal" and eligible applicants are identified.</p> <p>Per the requirement in Volume 1 Part E "Project Contributors" (pg. 27), it is the responsibility of the Offeror to identify all anticipated Project Participants and define the organizational structure, functional roles, and tasks.</p>

11	Can you clarify the export restrictions of this program – will it be ITAR or EAR?	Export Control Law requirements are addressed in the amended BAA Section II.C.1 (pg 21) and Section II.F.2.b.iii (pg. 57). Applicants are responsible for addressing any and all export control restrictions in their proposal
12	I tried to logon to IDEAS and this BAA wasn't pulling into the site –	The BAA will be added to IDEAS the week of April 10th.
13	Section II.D.2.d states "proposals shall be submitted electronically through the IARPA Distribution and Evaluation System (IDEAS);" however, as of today (04/02/2023) the REASON BAA is not available as a selection on the IDEAS website (at least not for offerors attempting to register with IDEAS for the first time). Can you provide a timeline on when the REASON BAA will be selectable in IDEAS?	The BAA will be added to IDEAS the week of April 10th.
14	The proposal requires submission through IDEAS, however as of 4/3/23, we were unable to register with the appropriate BAA. When will registration be available?	The BAA will be added to IDEAS the week of April 10th.
15	The BAA states to submit proposals electronically through the IARPA Distribution and Evaluation System (IDEAS): https://iarpa-ideas.gov . To get into the site, you need to select a BAA, yet there are no BAA's listed. Will the site be updated soon to include the REASON BAA?	The BAA will be added to IDEAS the week of April 10th.
16	For Volume 1 Technical and Management Proposal, is the requirement to submit one file with Sections 1, 2, 3, and 4 in it?	Per the BAA Amendment 1, the instructions in Section II.D.2.b Content and Form of Application Submission (pg 25), "All information pertaining to a volume shall be contained within that volume."
17	For Volume 1 Technical and Management Proposal, are Attachments 1-10 required to be in the same file as the main volume or in separate files, with one file for each attachment?	Per the BAA Amendment 1, the instructions in Section II.D.2.b Content and Form of Application Submission (pg 25), "All information pertaining to a volume shall be contained within that volume."
18	Should Attachment 8 be submitted as a separate PowerPoint file, or does the Government require that images of the slides/charts be placed in the Volume I file?	Per the BAA Amendment 1, the instructions in Section II.D.2.b Content and Form of Application Submission (pg 25), "All information pertaining to a volume shall be contained within that volume."
19	Our institution is a public university, and university travel policy dictates that we apply State of Arizona travel rates when budgeting. Would it be allowable for us to use State travel rates, or would we need to use JTR rates (as mentioned on p. 38 of the BAA)?	Per the Amended BAA, see Section II.D. Volume II - Cost Proposal titled, "Cost Proposal (No Page Limit)" (pg. 40)
20	Is there a cap limit of requested funds?	Per the Amended BAA, see Section II.D. Volume II - Cost Proposal titled, "Reasoning for Submitting a Strong Cost Proposal" (pg. 39-40)
21	Page 15, Figure 1 Schedule of Key Milestones and Deliverables. The schedule graphic and chart indicate a continuous flow between Phase1 and Phase 2. However, no 'waypoint' is given for a GO/NOGO for phase 2 – A break between the two phases a more than a month may have an impact upon cost projections for the Phase 2. Is there an anticipated gap and data on how long might it be? (Later sections indicate no gap – but if true, when is the go/no-go date?)	No, there will not be a gap between Phase 1 and Phase 2.

22	Will the Government exempt this Gantt chart from the page limits of Volume 1 to enable detailed information to be provided?	No, this is not compliant with the BAA instructions in Section II.D.2.b (pg 24-25)
23	Will the Government permit that Gantt charts use 11"x17" paper (foldouts) in order to present the detailed information in the required 10-point font size or exempt the Gantt charts from the 10-point font requirement?	No, this is not compliant with the BAA instructions in Section II.D.2.b (pg 24-25)
24	Why is Volume 1's page limit 21-pages? Volume 1 consists of the Summary of the Proposal and a Detailed proposal. The Summary Proposal shall not exceed 5 pages, the Detailed Proposal shall not to exceed 15 pages.	See updates in amended BAA published on 5 April 2023
25	On page 24, the BAA states that "Except for the cover sheet, transmittal letter, table of contents (optional), and the required attachments stated in the BAA, Volume 1 shall not exceed 21 pages." However, further instructions say that Section 2 should not exceed 5 pages, and Section 3 is up to 15 pages (totaling 20 pages). Which number is correct?	See updates in amended BAA published on 5 April 2023
26	The BAA example for the three chart summary contains examples for chart one and chart three, but not for chart two. Please provide an example of what is expected in chart two.	See updates in amended BAA published on 5 April 2023
27	With regards to the 3 Chart Summary, page 32 of the instructions state, "The format for the summary slides is included in Section II.H.8 to this BAA." The slides provided in Section II.H.8 (page 63) are missing Chart 2. Are there specific instructions for Chart 2?	See updates in amended BAA published on 5 April 2023
28	What information is required in Chart 2?	See updates in amended BAA published on 5 April 2023
29	Pg. 63 states that it is an "Example of Three Chart Summary of the Proposal;" however, there are only two charts (Chart 1 and Chart 3) shown. Can IARPA provide an outline of what information should be included in Chart 2?	See updates in amended BAA published on 5 April 2023
30	Will the Government permit colored backgrounds behind text in tables and graphics to enable color coding to organize information presented?	Yes, as long as it is compliant with amended BAA Section II.D.2.b (pg 24-25)
31	Is Section 1 the correct place for the Table of Contents, after the Cover Sheet and Transmittal Letter?	Yes, as per the instructions in amended BAA Section II.D.2.b on page 26 titled "Volume 1: Technical and Management Proposal"
32	The BAA specifies, "F. Attachment 6: Consultant Commitment Letters" but there is no additional information about this requirement. Are commitment letters for contingent hires and potential consultants [sic] to be included in Attachment 6?	Offerors can use their own format for the Consultant Commitment Letters. Commitment letters are only needed for Consultants.
33	Is there a specified format or required information for the consultant commitment letters, or may offerors use their own formats?	Offerors may use their own formats.
34	Regarding Section II.D.2, "Volume 2 – Cost Proposal" (page 32), in addition to submitting the numerical cost and pricing data using Microsoft Excel for the "Volume 2, Section 2, Estimated Cost Breakdown," can a PDF cost narrative be submitted to include information such as, a basis of estimate, direct/indirect rate information, proposal assumptions, etc.?	See updates in amended BAA published on 5 April 2023

35	The cost proposal is to be provided in Microsoft Excel in the format provided in Section II.H. Can an Excel file of this template be found online?	See updates in amended BAA published on 5 April 2023. The Government will not provide an Excel file for this template.
36	We must break down the major program tasks by fiscal year. Is this government fiscal year, offeror fiscal year, or both?	Government fiscal year.
37	The offeror must detail the cost breakdown to the Contracting Officer. Is this done with a separate uploaded PDF budget justification document or done within the Excel file alone?	See updates in amended BAA published on 5 April 2023
38	Starting at page 40, the BAA has a section c. Preparing an Application which provides instructions which appear to be similar-but-different to the instructions starting on p.22. The former appear to be instructions for a proposal to Army whereas the latter appear to be instructions for a proposal to IARPA. Can you clarify the relationship between these instructions.	See updates in amended BAA published on 5 April 2023
39	<p>I have read through the guidelines W911NF-23-S-0007. Can you clarify the required documents needed? To me, it read that there were different requirements if applying through IDEAS (pages 23-40) vs. an email submission (pages 40-47).</p> <p>For example, it looks like some documents such as Biographical Sketches and Current and Pending Supports were needed in the email submission, but were not asked in the IDEAS submission.</p> <p>Similarly, the content of the 15-page technical document seems to differ between those 2 sections.</p> <p>I wanted to make sure I was informing the faculty with accuracy what was required, so any clarity you can provide is greatly appreciated.</p>	See updates in amended BAA published on 5 April 2023
40	Which set of instructions takes precedence?	See updates in amended BAA published on 5 April 2023
41	If the additional information specified in pages 40-46 are required, where in the proposal must they be provided, and will the Government exempt them from page limits?	See updates in amended BAA published on 5 April 2023

42	<p>Section II.D.2.b. “Proposal Format” (page 23) indicates a compliant proposal should contain: Volume 1 – Technical and Management Proposal, with Sections 1-4; and Volume 2 – Cost Proposal, with Sections 1 and 2. Section II.D.2.c. “Preparing an Application” (pages 40-46) indicates the following forms should be included in the proposal if submitted via e-mail: Cover Sheet; Research and Related Senior/Key Person Profile (Expanded) form; Research and Related Personal Data form; Table of Contents; Statement of Disclosure Preference (Form 52 or 52A); Research and Related Other Project Information; Project Abstract; Technical Proposal (Project Description); Biographical Sketches; Bibliography; Current and Pending Support; Facilities, Equipment and Other Resources; and Cost Proposal (including DD Form 1861). However, Section II.D.2.d “Submission of Complete Research Proposals” (page 46) indicates “Proposals shall be submitted electronically through the IARPA Distribution and Evaluation System (IDEAS).” If proposals must be submitted through IDEAS, are the forms listed under Section II.D.2.c “Preparing an Application” required for a compliant proposal? And if required, should they be provided as stand-alone documents, or should they be incorporated into the technical/cost volumes?</p>	See updates in amended BAA published on 5 April 2023
43	<p>Section II.D.2.c (pages 40-46) of the BAA outlines a different proposal format. It looks like these instructions apply to “all proposals submitted via email.” However, later on the BAA states “Volume 1 submitted by any means other than IDEAS (e.g., hand-carried, postal service, commercial carrier and email) shall not be considered.” Is IDEA [<i>sic</i>] the only mean for proposal submission?</p>	See updates in amended BAA published on 5 April 2023
44	<p>Are the attachments mentioned in the “only submission via email” (BAA pages 40-46) section required for electronic submission as well (i.e. biosketch, current and pending support, facilities, Abstract, DD Form 1861)?</p>	See updates in amended BAA published on 5 April 2023

45	<p>There are two structures for the technical and management volume stated in the REASON BAA. These two structures are found in Section D.2.b, PP 23-29 and D.2.c, PP 42-45. We are submitting via IDEAS, so we are planning to follow the proposal outline on PP 23-29. The proposal outline on PP 42-45 is for entities submitting via email per section c. Preparing an Application, starting on page 40. Is this interpretation and application of the instructions in the BAA correct?</p>	See updates in amended BAA published on 5 April 2023
46	<p>Section II.D.2.c (Preparing an Application), pg. 40, states "This format applies to all proposals submitted via email" while Section II.D.2.d states "proposals shall be submitted electronically through the IARPA Distribution and Evaluation System (IDEAS)."</p> <p>Section II.D.2.c describes elements of a proposal that differ to some extent from the the proposal elements described elsewhere in the BAA (for example, the Cover Sheet described on pg. 40 lists elements not included in the Cover Sheet outline on pg. 56 and the Technical Proposal (starting on Pg. 42) description includes additional sections and elements not found in the Volume 1: Technical and Management Proposal outline (starting on Pg. 26).</p> <p>Do the items described in Section II.D.2.c NOT apply to proposals submitted in IDEAS (since Section II.D.2.c states "this format applies to all proposals submitted via email")?</p> <p>Or, if the items in Section II.D.2.c DO apply to proposals submitted in IDEAS, can IARPA provide guidance to offerors as to how to handle discrepancies between proposal items described in Section II.D.2.c and those described elsewhere in the BAA?</p>	See updates in amended BAA published on 5 April 2023
47	<p>Is there an Excel budget template that the program would like us to use? I understand that the format should be similar to the example in II.H (p. 67 of the BAA) but wondered if there was an existing template that would provide the level of detail needed by IARPA's cost analysts.</p>	See updates in amended BAA published on 5 April 2023. The Government will not provide an Excel file for this template.
48	<p>The BAA provides the clause FAR 52.209-11: Representation by Corporations Regarding Delinquent Tax Liability or a Felony Conviction under any Federal Law (Feb 2016). If offerors are required to complete this representation, may it be placed in Volume 2 Cost Proposal?</p>	Per the Amended BAA Section II.D.2.b (pg. 26), the Transmittal Letter shall include, "...any non-negotiable conditions on which the offer is based..."; and any information included in other parts of the proposal (ex. Volume 2 Cost Proposal), should be consistent with the information in the Transmittal Letter.

49	The BAA specifies, "J. Attachment 10: Privacy Plan" but does not provide a sample or a list of required information. (Page 11, paragraph b also references Performer-provided data.) Will the Government please provide a sample for this documentation or a list of information required in the attachment?	Contractor format is acceptable. Offeror's Privacy Plans shall meet the requirements in Section II.A.5.b on page 11 on the BAA amendment.
50	Is the Privacy Plan in Attachment 10 (BAA page 31) or Attachment 6 (BAA page 11)?	The Government acknowledges the administrative error, and clarifies that the Privacy Plan is attachment 10.
51	If the schedule and milestones document is provided in a separate Excel file, will it be exempted from the Volume 1 page limitations?	No, this is not compliant with the amended BAA Section II.D.2.b on pages 24-25.
52	If the schedule and milestones document is provided in a separate Excel file, will it be exempt from the requirement for 8.5"x11" page size and 10-point font?	No, this is not compliant with the amended BAA Section II.D.2.b on pages 24-25.
53	<p>Additionally, on page 25 of the BAA, the instructions state:</p> <p>a. Summarize total level of effort by labor category and technical discipline (i.e., research scientist/chemist/physicist/engineer/administrative, etc.) and affiliation (prime/ subcontractor/consultant). Key Personnel shall be identified by name. Provide a brief description of the qualifications for each labor category (i.e., education, certifications, years of experience, etc.)</p> <p>b. Summarize level of effort by labor category and technical discipline for each major task.</p> <p>How should we capture this information? The "Example of Team Organization Table" on page 59 does not capture this information.</p>	It is at the discretion of the Offeror on how to present the information in the Technical Resource Summary in Section II.D.2.b (pg 27 of the Amended BAA).
54	A sample is provided "Example of Team Organization Table" but this table is not mentioned in the BAA. Is this "Team Organization Table" in H.4 on page 59 one and the same as the "organization chart" required for Volume 1, Section 2, paragraph E on page 29?	The example in II.H.4 is an example of how to provide the required information in Section II.D.2.b under Project Contributors (pg 27 of the amended BAA).

55	<p>Section 3: Detailed Proposal Information includes the following sub-sections:</p> <ul style="list-style-type: none"> A. Statement of Work B. Objectives, Scientific Relevance, Technical Approach, and Expected Significance C. State of the Art D. Data Sources E. Deliverables <p>A previous IARPA BAA also had these sub-sections:</p> <ul style="list-style-type: none"> F. Cost, schedule, milestones G. Offeror's previous accomplishments H. Facilities I. Detailed Management Plan J. Resource Share K. Other Agencies or Parties <p>Can you clarify whether sections F-K are intentionally not included, and how proposals should handle content which might have been included in those sections.</p>	See updates in amended BAA published on 5 April 2023
56	<p>Could the Government please confirm that these three Task Areas must be addressed in "Section 3 Detailed Proposal Information" within Volume 1 Technical & Management Proposal?</p>	Yes
57	<p>The BAA specifies, "Proposals shall explicitly address all the following:</p> <ul style="list-style-type: none"> •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." Does the Government require that these be addressed in Section 3? May they be addressed anywhere within the section, or does the Government require that they be addressed together as in information set? 	Per the Amended BAA Section II.D.2.b (pg. 24), "To facilitate the evaluation of the proposal, the government encourages the offerors to submit proposals which: are clear and concise; limited to essential matters sufficient to demonstrate a complete understanding of the Government's requirements; include sufficient detail for effective evaluation; and provide convincing rationale to address how the offeror intends to meet these requirements and objectives, rather than simply rephrasing or restating the Government's requirements and objectives."

58	Is there any requirement (or something similar) that Principal Investigators and/or key personnel on this program must commit to (3) months of effort per year?	See updates in amended BAA published on 5 April 2023, including Section 3 Part I Detailed Management Plan (pg. 31 -32)
59	Could you please confirm how this proposal should be submitted? On page 40 of the BAA, it discusses email submissions, but then on page 46, it states that the proposal must be submitted via IDEAS.	See updates in amended BAA published on 5 April 2023
60	Can we add information outside of the corpus that is provided by the government?	Per BAA (Amendment 1 of 5 APR 23) Section II.A.5.b, "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." You may use additional information for training data. However, Performer system responses to Challenge Problems may only cite items in the corpus. Per BAA Section II.A.7, "The Performer systems produce a set of up to 8 ordered evidence items found in the corpus as the output for the Challenge Problem."
61	Can we use external data on disparate geographic areas such that our processes can make analogous comparisons with this other data?	Per BAA (Amendment 1 of 5 APR 23) Section II.A.5.b page 11, "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." You may use additional information for training data. However, Performer system responses to Challenge Problems may only cite items in the corpus. Per BAA Section II.A.7, "The Performer systems produce a set of up to 8 ordered evidence items found in the corpus as the output for the Challenge Problem."
62	Under TA1, the BAA directs proposers to identify "additional evidence contained in the corpus but not mentioned in the draft report." The technology we are proposing can also identify possible evidence not in the corpus, but whose existence would support or weaken the claims of the report. E.g., "Your claim in paragraph 3 would be much stronger if you can find evidence showing that XYZ." Information of this sort is valuable for preparing prioritized information requests (PIRs). Is such potential evidence desired in a response to the TA1 tests, or should evidence listed be restricted to that in the corpus?	Performer system responses to Challenge Problems may only cite items in the corpus. See BAA (Amendment 1 of 5 APR 23) Section II.A.7 page 13: "The Performer systems produce a set of up to 8 ordered evidence items found in the corpus as the output for the Challenge Problem."
63	Can the corpus be supplemented with additional training data, such as scholarly datasets?	BAA (Amendment 1 of 5 APR 23) Section II.A.5.b on page 11: "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." You may use additional information for training data. However, Performer system responses to Challenge Problems may only cite items in the corpus. See BAA Section II.A.7 on page 13: "The Performer systems produce a set of up to 8 ordered evidence items found in the corpus as the output for the Challenge Problem."

64	<p>Additional datasets provided by the performer:</p> <p>(a) are there any restrictions on data involving US persons since this is an IC program? Specifically we have had experience with other IC agencies restricting use of Twitter data from the US because of involvement with US Persons. Please clarify in detail such issues for REASON. In particular, can public domain data involving US persons (e.g. news articles and Twitter) be supplied by the performer for this program?</p> <p>(b) Are datasets provided by the performer required to be made available to other performers?</p> <p>(c) Are datasets provided by the performer required to be put in the public domain?</p> <p>(d) Can the performer work with sensitive but unclassified datasets for testing purposes? One example might be written data from a domestic law enforcement agency (please note also the connection to US persons here).</p>	<p>Question a: Requirements for performer-provided data are provided in BAA (Amendment 1 of 5 APR 23) Section II.A.5.b, including the requirement for a Privacy Plan. (Page 11).</p> <p>Question b: Performers are not required to share data with other Performers.</p> <p>Question c: See the requirements on the IARPA web site for making data available for publicly funded research <https://www.iarpa.gov/research-programs/public-access-to-iarpa-research>. Other guidance will provided at Program Kickoff.</p> <p>Question d: BAA (Amendment 1 of 5 APR 23) Section II.A.4 page 10 states "All Performer research will be <i>strictly</i> unclassified" (italics added for emphasis).</p>
65	<p>Can the performer provide datasets in a structured format for intermediate step internal testing of algorithms? One example might be data in the form of knowledge graphs, if knowledge graphs are being used as a step in the algorithm design.</p>	<p>BAA (Amendment 1 of 5 APR 23) Section II.A.5.b page 11: "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."</p>
66	<p>Can data developed by the performer (5b, page 11) be used as part of the model being evaluated (e.g. in a non-parametric approach)?</p>	<p>Per the BAA (Amendment 1 of 5 APR 23) Section II.A.5.b page 11: "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."</p> <p>In addition, per the BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16: "the entirety of a Performer's system for TA1, TA2, and TA3 REASON Comment Quality evaluations, including pre- and post-processing, must be included within the delivered software container."</p>
67	<p>Can we get an example of "apparent" vs. "real" strengths and weaknesses as described in section 2 at the top of page 7.</p>	<p>One type of difference between apparent vs. real strength in reasoning is the difference between being well-written and being well-reasoned. Another is the difference between being persuasive and being well-reasoned.</p>
68	<p>Regarding page 7, could you clarify your thinking around "distinguish between apparent and real strengths"</p>	<p>One type of difference between apparent vs. real strength in reasoning is the difference between being well-written and being well-reasoned. Another is the difference between being persuasive and being well-reasoned.</p>
69	<p>"A successful system must distinguish between apparent and real strengths and weaknesses ..." --</p> <p>Can the government clarify how it will objectively define and delineate between 'apparent' and 'real' strengths and weaknesses?</p>	<p>One type of difference between apparent vs. real strength in reasoning is the difference between being well-written and being well-reasoned. Another is the difference between being persuasive and being well-reasoned.</p>

70	Will the draft analytic reports have citations to the evidence that is used in the report?	This depends on the TA being evaluated. TA1 and TA3 Reason Comment Quality Challenge Problems will include some citations for evidence. Per the BAA (Amendment 1 of 5 APR 23) Section II.A.6 page 12: "The format for Challenge Problems will vary across the TAs and will be specified in the Phase 1 T&E Plan at Program Kickoff."
71	Regarding T&E, will teams write analytic reports or will individuals? How many reports will be written on each problem? Given the need for statistical power to compare the AI conditions from the control conditions, it is not clear how T&E will be able to have sufficient to compare the systems from the control.	The T&E Plan will be shared with Performers at, or shortly after, kickoff and will include these details. Regarding the TA3 Final Exam in each phase, described in BAA (Amendment 1 of 5 APR 23) Section II.A.3 on Page 8, the work will be performed by individuals, not teams.
72	Would it be possible to receive a challenge example made of an analytic question, a draft report and a related corpus of surrogate source documents?	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.6, page 12: "For TA1, TA2, and TA3 REASON Comment Quality evaluations, T&E will provide 5 Practice Challenge Problems as examples prior to the first cycle for that type of evaluation. The Practice Challenge Problems will include the inputs to Performer systems, an emulated Performer system solution, and annotation of that solution according to the evaluation procedure in the T&E plan."
73	Can we know in advance the domains (e.g., chemistry, pharmacy ... academic research, computer science ...) that T&E will be performed on in order to obtain the right collection of training data?	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.3 page 8: "The analytic questions will be drawn from a wide variety of topic areas, including political, military, social, economic, environmental, or diplomatic topics."
74	The BAA states "For each TA, an evaluation cycle will consist of approximately 25 Challenge Problems developed by T&E. The format for Challenge Problems will vary across the TAs and will be specified in the Phase 1 T&E Plan at Program Kickoff. Within an evaluation cycle each Performer system will receive the same Challenge Problems. Challenge Problems will differ across evaluation cycles. For TA1, TA2, and TA3 REASON Comment Quality evaluations, T&E will provide 5 Practice Challenge Problems as examples prior to the first cycle for that type of evaluation. The Practice Challenge Problems will include the inputs to Performer systems, an emulated Performer system solution, and annotation of that solution according to the evaluation procedure in the T&E plan." It is not clear how many of the 25 Challenge Problems per T&E cycle will be assigned to each of the TAs (all 25 or some subset). Can you provide an approximate breakdown of the number of problems per each TA per cycle?	The description of evaluation cycles is found in the BAA (Amendment 1 of 5 APR 23) Section II.A.6 page 12: "For each TA, an evaluation cycle will consist of approximately 25 Challenge Problems developed by T&E."
75	On page 17, the BAA states "Offerors must specify the runtime resources and services required for their delivered software in terms equivalent to a configuration on either Amazon Web Services, Microsoft Azure, or Google Cloud." We understand that this will help the Government and T&E to plan for evaluations on unclassified material. However, is this consistent with the requirement that T&E will train and test on classified data, not just on unclassified data?	Per the BAA (Amendment 1 of 5 APR 23) Section II.8.c page 17: "The API will be functionally identical for unclassified and classified testing." The classified testbed will be designed to be very similar to the unclassified testbed so the resource descriptions and limits will be germane for classified testing.

76	<p>What kind of resources will the Docker container have access to? GPU? RAM? Disk space? ports?</p>	<p>These details will be provided to Performers at Program Kickoff. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 17: "Offerors must specify the runtime resources and services required for their delivered software in terms equivalent to a configuration on either Amazon Web Services, Microsoft Azure, or Google Cloud." This section also states: "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."</p>
77	<p>At what point in the program must "Offeror... specify the runtime resources and services required for their delivered software in terms equivalent to a configuration on [a cloud provider]"? What is the expected resourcing available to government to train the system high side? Is there a preferred cloud environment and/or other HW/SW components?</p>	<p>Per the BAA (Amendment 1 of 5 APR 23) Section II.8.c page 17: "The API will be functionally identical for unclassified and classified testing." The offeror must specify these requirements in their Technical Proposals. The classified testbed will be designed to be very similar to the unclassified testbed so the resource descriptions and limits will be germane for classified testing. This section also states: "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."</p>
78	<p>Can the government provide some info on the software, computing environment, and infrastructure used for the T&E evaluation?</p>	<p>These details will be provided to Performers at Program Kickoff. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 17: "Offerors must specify the runtime resources and services required for their delivered software in terms equivalent to a configuration on either Amazon Web Services, Microsoft Azure, or Google Cloud." This section also states: "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."</p>
79	<p>What budget / resource limitations are expected for the operation of the REASON system? For example, inference requiring use of large language models may require many GPUs.</p>	<p>These details will be provided to Performers at Program Kickoff. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 17: "Offerors must specify the runtime resources and services required for their delivered software in terms equivalent to a configuration on either Amazon Web Services, Microsoft Azure, or Google Cloud." This section also states: "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."</p>
80	<p>For the unclassified data mentioned on page 10 (5.a.) would we expect only news articles related to the intelligence reports (e.g., will it be restricted to the same geo as the intel reports) and will such information only be intelligence-relevant (e.g. will it include sports or arts sections of these media sources)?</p>	<p>The unclassified data will have broad topical and geographical coverage. The Government does not currently plan to filter out portions of the news items. BAA (Amendment 1 of 5 APR 23) Section II.A.5.a page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets."</p>

81	Section 5a states: "At the beginning of Phase 1 this will be a corpus of unclassified news articles and analytic reports" (p. 10). Are the "Analytic Reports," referenced standard analytic reports by IC analysts, or are they the 5 sample draft reports prepared by T&E?	The section referred to can be found in BAA (Amendment 1 of 5 APR 23) Section II.A.5.a. on pages 10 and 11. Details for the corpus will be provided at Program Kickoff. The intention is that the analytic reports in the corpus will consist of open source items that are qualitatively similar to analytic reports. With the possible exception of declassified IC reports, they will not in general be prepared by IC analysts. The 5 sample draft reports from T&E are included as part of worked Challenge Problems that provide Performers with example problems and expected responses.
82	We understand that the government intends for the T&E team to re-train and execute the applications on classified data. Can the government confirm that the size and content of the classified training sets will match the unclassified corpus that performer teams will be provided?	The Government will not provide information about the size or content of the classified training sets.
83	Best approaches may vary based on number of documents, and number of documents between surrogate and T&E datasets. Approximately how many documents will be in the surrogate data, and how many will be included in the T&E data?	The surrogate and T&E data sets are the same. As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.5.a page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets." The Government will not provide information about the size or content of the classified training sets.
84	What is the anticipated ratio of document amounts between the unclassified set and the classified set?	The Government will not provide information about the size or content of the classified training sets.
85	What is the expected data size (# of articles, bytes) of the source document corpus?	As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.5.a, page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets."
86	Can the government elaborate on the source(s) and nature of the data? Is it from a commercial provider?	As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.5.a, Page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets."
87	Please provide a minimum of 10 sample draft analytic reports and news articles for review.	The Government will provide that information at Program Kickoff.
88	Please provide a list of the 20 media outlets referenced on Page 10 of BAA?	The Government will provide that information at Program Kickoff.
89	Can we have prior to proposal submission one example of declassified data that an analyst has summarized, with a declassified version of the summary?	The Government will not provide that information.
90	Supporting and refuting evidence is to be drawn from a "corpus of source documents"; will this be a single corpus or many different corpora that vary based upon the contents of the draft analytic report(s)? How large will said corpora be with the diverse set of content? (news reports, analytic reports, analysis and opinion articles and reference materials) For example: 5TB, 20TB, etc. or 1MM pages, 20 MM pages, etc.	As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.5.a, Page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets."

91	What are the expected sizes of the unclassified and classified corpora? Will this be on the order of MB? GBs? TBs?	As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.5.a, Page 10: "Performers may assume that this corpus of data will include at least 25 years of output from at least 20 major reputable media outlets."
92	What type of metadata will be included with the unclassified and classified corpora?	This detail will be available to Performers at Program Kickoff
93	Will the date and time of publication for each document be included with the unclassified and classified corpora?	This detail will be available to Performers at Program Kickoff
94	What is the expected length of the reports/documents? Will this differ substantially between the train and test?	This detail will be available to Performers at Program Kickoff
95	For each document in the corpus, what level of pre-processing is required by the teams? Do we need to parse special types of files (e.g. HTML), or will the documents be provided as TXT?	This detail will be available to Performers at Program Kickoff
96	How will the organizers make the corpus of documents available to the system? Will it be provided as an S3 bucket or similar?	This detail will be available to Performers at Program Kickoff
97	Will the organizers provide the domains of each of the documents (e.g. news, analyst reports, etc) as part of this metadata?	This detail will be available to Performers at Program Kickoff
98	Will documents have hyperlinks or other document identifiers that reference other documents in the corpus?	This detail will be available to Performers at Program Kickoff
99	Many of the (declassified) analytic reports my team has reviewed do not present the evidence on which they rely: they read more like summaries of inferences drawn from undisclosed evidence than like arguments supported by evaluable reasoning. Will the analytic products shared with REASON performers include descriptions of pertinent "raw" evidence, or will they be more like the summaries I have seen so far? If the evidence that analysts relied on isn't provided in the reports accessible to performers, will it be made available in some other way, and if so, how will it be indexed to the reports?	Draft analytic reports included in Challenge Problems will be constructed to resemble draft forms of IC analytic reports and they will generally cite the evidence they rely on. Further details about surrogate or draft analytic reports in the corpus will be provided at Program Kickoff.
100	The Proposal requires that the REASON technology fit into the existing intelligence analysts' workflow, can you please provide a diagram of the current workflow? We would also request the current application(s) which are used by the intelligence analyst; We understand they use word, but what other applications/databases do they connect to gather evidence or assist with the writing of the draft?	An outline of the workflow was presented at REASON Proposers Day, which can be found here: https://www.iarpa.gov/research-programs/reason
101	Can we leverage large language models' APIs, to access chatGPT or GPT-4?	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16: "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.
102	Is the software allowed to connect to external APIs, such as the S2ORC Graph API, the GPT APT, the Microsoft Cognitive API, and the DBPedia API?	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16: "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.

103	Are we allowed to leverage large language models such as chatGPT, GPT 4 as part of our model? A general question is whether we are allowed to incorporate any models from commercial world.	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16: "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.
104	During the performance period, can we submit program evaluation data to third-party API services such as ChatGPT or GPT-4?	As stated in BAA (Amendment 1 of 5 APR 23) II.A.8.b page 16: "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed. If the question is about submitting T&E data to a third-party then guidance will be provided at or after Program Kickoff.
105	Presumably, there are many relevant content repositories that analysts use to provide reasons for linking data with conclusions. E.g., known locations of terrorist groups or individuals. Will there be unclassified analogues of these systems made available to us? If so, please provide a list of any and all systems required for Integration during this program.	REASON T&E will not rely on such databases or their surrogates. All evidence will be found in the corpus of documents.
106	Is the Government referring to IRB requirements for human subjects? If not, will the Government clarify?	The Government is referring to the requirements in BAA (Amendment 1 of 5 APR 23) Section II.A.13.
107	Human subject use: The discussion about human subjects in the BAA suggests that such people might be used for internal evaluation of methods by the performer. Was this your intent or would there be other reasons to include human subjects in the research.	Per the BAA (Amendment 1 of 5 APR 23) Section II.A.13 page 19: "Performer human subjects research for REASON is encouraged but not required." and "Performers are not responsible for obtaining IRB approval for official T&E evaluation events. For these events the T&E organization will obtain the necessary approvals." Each offeror will need to determine whether to utilize human subjects or how to use human subjects in internal development and testing.
108	Can we use open-sourced large language models (e.g., GPT-J or LLaMA) or their fine-tuned versions (on our data or the project-related data) as part of our system?	Yes, provided that the offeror has appropriate IP rights and can include the models in their containerized delivery. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16, "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.
109	Can we incorporate open-sourced models (such as the one from Meta)?	Yes, provided that the offeror has appropriate IP rights and can include the models in their containerized delivery. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16, "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.
110	It is mentioned that non-textual data is out of scope, but that it is permissible to use image captions to locate and retrieve non-textual items. This seems somewhat contradictory; for what purpose are we retrieving non-textual items of analyzing them is out of scope?	Performer systems in TA3 will be permitted to present non-textual information as part of a suggestion to the analyst.
111	Should failure to meet ICD 206 standards for citing evidence be scored as a weakness?	Detailed submission formats in response to challenge problems will be provided to Performers at, or shortly after, Program Kickoff. However, the TA1 focus is on identifying the textual content that provides evidence.

112	<p>The BAA excludes processing of graphical information from the scope of the program. However, ICD 203 "Analytic Standards" Section D.6.e.(9) requires IC analytic products to</p> <p>Incorporate effective visual information where appropriate: Analytic products should incorporate visual information to clarify an analytic message and to complement or enhance the presentation of data and analysis. In particular, visual presentations should be used when information or concepts (e.g., spatial or temporal relationships) can be conveyed better in graphic form (e.g., tables, flow charts, images) than in written text. Visual information may range from plain presentation of intelligence information to interactive displays for complex information and analytic concepts. All of the content in an analytic product may be presented visually. Visual information should always be clear and pertinent to the product's subject. Analytic content in visual information should also adhere to other analytic tradecraft standards.</p> <p>Ordinarily, diagrams would be very useful to express the structure of an argument with precision. Are we to expect that the analytic products given to us to analyze will deviate from this standard, and that all of the products given to us will define their argumentation entirely verbally? Or should the restriction on using graphical information be reconsidered?</p>	<p>The cited standard is not one of the 6 which will be evaluated in the TA3 Final Exam. According to the BAA (Amendment 1 of 5 APR 23) Section II.A.7, page 14: "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." In BAA (Amendment 1 of 5 APR 23) Section II.A.4 on page 10, processing of non-textual inputs is out of scope for REASON. Please note that the requirements for TA3 comments to analysts (BAA Section II.A.2, pages 6 and 7) permit the inclusion of graphics with those comments.</p>
113	<p>Can the Word plug-in communicate with the Docker container through a Restful API? Put another way, does the user interface need to be containerized?</p>	<p>The user interface does not need to be containerized. Full details will be provided after Program Kickoff.</p>
114	<p>Will performers be able to demonstrate their capabilities with commercially-available approaches that are not available in the classified environment, or do performers have to demonstrate capabilities keeping in mind constraints in the classified environment?</p>	<p>As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8. page 16: "the entirety of a Performer's system ...must be included within the delivered software container." It may not access resources outside of the testbed.</p>
115	<p>Last paragraph on p. 16 says: 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. Can we get specifics on what are actually required in advance?</p>	<p>As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b, page 16: "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)."</p>

116	Can we use commercial API or open-source models in the final software delivery?	Yes, provided that the offeror has appropriate IP rights and can include the models in their containerized delivery. As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b page 16, "the entirety of a Performer's system ... must be included within the delivered software container." It may not access resources outside of the testbed.
117	With the requirement that the software run in a Secure environment, even though the development in the unclass environment, are there any specific concerns we should have about acceptance into the classified environment? (e.g passing Fortify scans, not allowing certain languages, etc)	As stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8. page 16: "the entirety of a Performer's system ...must be included within the delivered software container." It may not access resources outside of the testbed. Further details will be provided at Program Kickoff.
118	Is assessing the reliability/trustworthiness of a source part of either TA2 or TA3?	The requirements for all TAs are given in BAA (Amendment 1 of 5 APR 23) Sections II.A.1 and II.A.2 on pages 6 and 7. Reliability/trustworthiness of a source may bear on the argumentation contained in a report. Assessments of this reliability/trustworthiness may be applicable to TA2 or TA3. We will discuss this in more detail with Performers at Program Kickoff.
119	TA1. Scoring: 4.b: How will the quality of the evidence be addressed? There will be many cases of relatedness which are using fallacious connections. In the BAA, categories such as public statements by leaders or analogous situations are mentioned. In argumentation theory and misinformation detection, many of them can be seen as very weak arguments not supported with actual facts. For example, many of the analogy arguments may be opinionated in nature, which also makes the evidence weak. In a subjective task such as this, it is essential to have a theoretical definition of a good evidence that the scorer will follow.	This Government interprets this question as addressing the metrics for TA1 in the BAA (Amendment 1 of 5 APR 23) Section II.A.7 pages 12-13. Relevance assessment in TA1 may consider quality of evidence. Evaluation of relevance will be performed by T&E, with details to be provided to Performers in the T&E plan provided at, or shortly after, Program Kickoff.
120	We thought we'd ask for clarification regarding the scope of the problem in TA1. When BAA talks about 8 pieces of evidence to be recommended in TA1, it was not entirely clear to us whether that means 8 pieces of evidence for the entire report or 8 pieces of evidence for each part of the report where evidence can be challenged/strengthened. We'd appreciate any insights.	The Government interprets that this question addresses the metrics for TA1 in the BAA (Amendment 1 of 5 APR 23) Section II.A.7 pages 12-13. The requirement is up to 8 items of additional evidence per report.
121	TA1: Analytic reports often contain more than one point. How will the scorer prioritise evidence in the cases when they refer to different points made in the report.	This Government interprets that this question addresses the metrics for TA1 in the BAA (Amendment 1 of 5 APR 23) Section II.A.7 pages 12-13. This distinction will be captured in the diversity consideration. Assuming equal relevance, a system that identifies evidence pertaining to multiple points will perform better than a system identifying evidence that only pertains to only one point.

122	TA1: The reasoning in the analytic reports often seems fallacious by nature, including the terms such as “we judge that” with no reference, “may”, “seems to”. In this light, it seems essential to organise the task areas 1 and 2 in a pipeline where the input to TA1 is the output of TA2. How are the organisers planning to address this dependency of the tasks?	TA1 and TA2 will be evaluated separately as depicted in BAA (Amendment 1 of 5 APR 23) Figure 1 on page 15. Offerors will need to determine, as part of their technical approach, components common to both TAs.
123	TA1. Scoring: 4.b: There can be many ways the evidence is relevant to the item. Different performers can potentially come up with their own underlying theory. How will this be addressed?	As the BAA(Amendment 1 of 5 APR 23) states in II.A.1 page 5: "evidence is information that supports or opposes a judgment." Evaluation of relevance will be performed by T&E, with details to be provided to Performers in the T&E plan provided at, or shortly after, Program Kickoff.
124	What constitutes an "item of additional evidence"? A statement? A document? Must it list the reason for the recommendation? Or will a human analyst attempt to infer why it supports or contradicts a conclusion? If the former, how will IARPA ensure objectivity in an inherently subjective assessment?	As the BAA(Amendment 1 of 5 APR 23) states in II.A.1 page 5: "evidence is information that supports or opposes a judgment." Additional evidence is evidence that is not already cited in the draft report. Evidence will be identified at a sub-document level. Further details will be specified in the T&E Plan at Program Kickoff. TA1 does not require reasons for recommending additional evidence. However, if an evidence item is the subject of a TA3 comment then the comment should provide an explanation.
125	For the TA1 module, what level of granularity is needed for identifying evidence (document level? section level? paragraph level? highlighting a span?)	Evidence will be identified at a sub-document level. Further details will be specified in the T&E Plan at Program Kickoff.
126	Are “categories” mentioned on page 13 (item 4 in section 7) provided by T&E or developed by performers?	The categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at or shortly after Program Kickoff.
127	Regarding page13 item (4c), how does the diversity of cited evidence feed into T&E assessment for TA1? How will evaluation balance diversity against relevance in scoring?	Relevance is primary. However, redundant evidence items are discounted in value. If two systems produce results with equal relevance then the one that has less redundancy among the first few items will score higher.
128	Regarding page 13 item (5), presumably the items discovered by T&E will be a list of the top 8 (and the formula is not the union of results, but an iterator over the sets of 8. Is that interpretation correct?	During Challenge Problem creation, T&E will manually search for and discover additional evidence items that are not included in the Challenge Problem. After prioritization by relevance and diversity, this becomes the first estimate of an ideal result set for the anDCG denominator. It is possible, however, that Performer systems will identify new evidence items that are more relevant or diverse from T&E's set. T&E will examine the results from all Performers in addition to their original set of evidence items. The best possible set of evidence items from the combination of T&E's set and Performer result sets will be used to determine the ideal result set for the denominator of the anDCG formula.

129	On page 13, how does the calculation of anDCG account for situations in which there are fewer than 8 items, i.e., on page 6, "When there is no non-redundant additional evidence, performer systems should report that"?	This will be addressed at Program Kickoff.
130	TA1. Scoring: 4.c: "The categories will be used to calculate the diversity of the cited evidence." - The BAA mentions that the evidence needs to be additionally classified. It also mentions, as an example, evidence of troop movements, public statements by leaders and previous examples of similar circumstances for the document about military invasions. It seems that these categories are very topic-specific. In this regard, will the set of topics and persuasion frames (called categories in BAA) be finite? The system that addresses framing classification needs to know what to look for in advance, and this needs to be a predefined set of classes per topic.	The categories listed in the BAA were provided only as examples. Categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at or shortly after Program Kickoff.
131	TA1. Scoring: 4.c: This is not clear what authors mean by scoring based on the category the evidence belongs to. Will some categories be of a higher priority than others and will this priority be given in advance?	The categories listed in the BAA were provided only as examples. Categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at, or shortly after, Program Kickoff.
132	TA1. Scoring: 4: each of the criteria in point 4 does not provide information regarding how the evidence should be prioritised given they comply with all three criteria: are the categories ranked? are the underlying relations ranked? what if the information is equally novel?	The categories listed in the BAA were provided only as examples. Categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at, or shortly after, Program Kickoff.
133	The last paragraph on page 7 says that TA-1 systems will be evaluated on their ability to find and rank-order evidence: Can the government clarify what objective criteria the rank-order evaluation is relative to?	Performer results will be compared against a computed ideal set of results. This ideal set will be determined by T&E manual search, augmented by the result sets returned by all Performer systems for the same Challenge Problem.
134	There are multiple formulations for computing Discounted Cumulative Gain in information retrieval applications. Specifically, will the numerator be the traditional formulation of the graded relative importance, or $2^{(\text{graded relative importance})} - 1$? We infer that the government's importance scoring is non-binary. Thus these formulations will be different. Can the government clarify the specific DCG formulation that will be used to score performer systems?	Exact evaluation details will be provided in the T&E Plan to be provided to Performers at, or shortly after, Program Kickoff.

135	<p>In Step 1, the T&E team will perform a manual search of the corpus for evidence. However, the government does not describe how the relative importance will be assigned. Is it purely based on information retrieval order? Or will there be a qualitative assessment of relevance? If the latter, will this process be repeated over multiple assessors to improve scientific credibility?</p>	<p>This will be a qualitative assessment of relevance by the Challenge Problem designers and raters. Details will be provided in the T&E Plan at Program Kickoff.</p>
136	<p>In step 4c, the government states it will calculate the category diversity of the performer-recommended evidence. However, the government does not describe how it will use this diversity metric. Is it purely for informational purposes, or will it inform performer evaluations?</p>	<p>The diversity factor will be used to discount redundant evidence in computing anDCG.</p>
137	<p>On Page 13 of BAA, point #4 of the scoring process states, “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 of the item? c)What category does the item belong to? ...” Does the program expect performers’ REASON tool to not only identify additional evidence, but also provide category labels for proposed additional evidence? If so, at what levels should these categories be given? Will the government provide the definition of categories?</p>	<p>The categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at, or shortly after, Program Kickoff.</p>
138	<p>Are the categories of evidence pre-defined as part of the domain or problem statement?</p>	<p>The categories will be used internally by T&E to determine the redundancy of listed evidence items and will be specific to the type of challenge problem. They will not be provided explicitly to Performers, nor will Performers be expected to output category labels. More details will be provided to Performers in the T&E Plan at, or shortly after, Program Kickoff.</p>
139	<p>On page 6, the BAA states “the system will identify up to eight strengths and weaknesses.” Does that mean up to 8 strengths and up to 8 weaknesses, or 8 total? If only 8 total, might it be better to just generate 8 weaknesses, since the strengths are already strong in the report?</p>	<p>A total of 8. TA2 evaluations will not distinguish between strengths and weaknesses for scoring. Providing prioritized comments to analysts is addressed in TA3.</p>
140	<p>On page 13, the calculation of recall for TA2 requires identification of *all* strengths and weaknesses in a report. Though some might be engineered in the drafting of the product, there will inevitably be (many) more strengths and weaknesses. Won't any measure of recall and therefore F1 suffer as a result? (This risk is typically mitigated in information retrieval evaluations by pooling rankings from many systems.)</p>	<p>The Government will not penalize Performers for identifying valid reasoning strengths and weaknesses even if they don't correspond to the T&E prior list. Full details will be provided to Performers at Program Kickoff.</p>

141	<p>TA2: it seems the scoring does not consider the correctness of identifying the argumentative parts of the text. The weaknesses and strengths addressed by different performers may refer to various arguments made in the report.</p>	<p>Evaluation criteria for TA2 are provided in BAA (Amendment 1 of 5 APR 23) section II.A.7, page 13. Further details, including submission formats and evaluation particulars, will be provided to Performers in the T&E Plan at, or after, Program Kickoff.</p>
142	<p>TA2: it seems very difficult to address the weaknesses or strengths of some arguments without seeing the evidence they refer to (which also leads to a connection to TA1). While we can derive certain signals from the structure, the arguments can be made based on fallacious links between the premises and conclusions. However, TA1 is slo inherently dependent on TA2, so both tasks rely on each other in some way.</p>	<p>TA2 Challenge Problems will include citations for evidence. TA1 and TA2 will be evaluated separately as described. Offerors will need to determine, as part of their technical approach, which software components are shared by TA1 and TA2.</p>
143	<p>The government states that performer TA-2 systems will be evaluated and scored on the Reasoning Explanation Quality of their results. However, the government will not define this metric until after the award and kickoff. This lack of upfront clarity in evaluation metrics inhibits industry's ability to propose credible technical approaches confidently.</p> <p>Is it possible for the government to delay proposal submissions until this metric can be defined?</p>	<p>Exact evaluation details will be provided in the T&E Plan to be provided to Performers at or shortly after Program Kickoff. The Government will not delay the proposal submission deadline.</p>
144	<p>The deductive-nomological model of explanation proposed by Hempel provides a set of criteria to start looking for acceptable explanations in science by connecting data to conclusion through scientific laws (or “law-like” statements). Is there a complementary model already being used by analysts in the task areas they work in? E.g., are there templates for the modus operandi of different terrorist groups? If so, please provide a list of any and all models currently in use or may require integration with our system during this program.</p>	<p>The Government will not provide that information.</p>
145	<p>TA3 guidelines note evaluation is based on assessments of the AI feedback’s ability to be ‘used effectively’ and then later notes that scoring criteria include being actionable, understandable, and useful. In another section it says TA3 evaluations are based on the comments being ‘correct, appropriate, and clear.’ Can we get any clarification on how TA3 outputs are being evaluated? And how those more abstract values are being operationalized?</p>	<p>Exact evaluation details will be provided in the T&E Plan to be provided to Performers at, or shortly after, Program Kickoff.</p>

146	<p>The government states that performer TA-3 systems will be evaluated and scored on the Reasoning Comment Quality Score of their results. However, the government will not define this metric until after the award and kickoff. This lack of upfront clarity in evaluation metrics inhibits industry's ability to propose credible technical approaches confidently.</p> <p>Is it possible for the government to delay proposal submissions until this metric can be defined?</p>	<p>Exact evaluation details will be provided in the T&E Plan to be provided to Performers at or shortly after Program Kickoff. The Government will not delay the proposal submission deadline.</p>
147	<p>While TA1 and TA2 are going to be graded on 8 recommended additional sources of evidence and 8 strength/weakness items respectively, the BAA did not specify a limit for how many comments TA3 may produce. Is there a limit or target number, or is this up to the performer and potentially dependent on the specifics of the draft report? If there is a limit or target, is that for the T&E evaluation of comment quality only, or also for the final exam?</p>	<p>Exact evaluation details will be provided in the T&E Plan to be provided to Performers at, or shortly after, Program Kickoff.</p>
148	<p>Can you provide references for previous empirical research that have used RCQ scores? Specifically, could you include the references that describe the study methodologies using a qualitative evaluation of correctness, appropriateness, and clarity of analytical commentary?</p>	<p>Guidelines for the RCQ evaluation will be included in the T&E Plan that will be available to Performers at Program Kickoff, or shortly thereafter.</p>
149	<p>Will performers receive the results of the government's evaluation of TA-1/2/3 systems to aid further iterations?</p> <p>Will cleared members of performer teams be permitted to see the classified results?</p>	<p>Performers will receive detailed results from evaluations using unclassified data. Performers will receive actionable summary level unclassified feedback from tests on classified data. As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.12 page 19: "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. Even if a Performer has cleared personnel, they will not receive additional classified feedback."</p>
150	<p>What are the procedures and methodology used for high side T&E in phase 2? How are results shared with performers and what is the granularity of the feedback?</p>	<p>Performers will receive detailed results from evaluations using unclassified data. Performers will receive actionable summary level unclassified feedback from tests on classified data. As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.12 page 19: "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. Even if a Performer has cleared personnel, they will not receive additional classified feedback."</p>

151	Will Performer have an opportunity to work with the classified Subject Matter Experts in SCIF if any problems arise setting up the containerized application? If so, what processes should be followed?	Performers will receive detailed results from evaluations using unclassified data. Performers will receive actionable summary level unclassified feedback from tests on classified data. As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.12 page 19: "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. Even if a Performer has cleared personnel, they will not receive additional classified feedback."
152	Will the T&E evaluation encompass training the model on T&E doc corpus (that could be different than the corpus provided to performers) or will T&E be only on performer-trained models used to produce TA1 output without retraining as part of the test process?	Unclassified tests will use the same corpus that is available to Performers. Classified tests will necessarily use a corpus that is unavailable to Performers. Section II.A.6 of the BAA (Amendment 1 of 5 APR 23) page 12 states "For classified testing, the T&E Team will retrain Performer systems to classified data as necessary, using scripts or processes provided by the Performers."
153	Is interaction with the intelligence analyst within scope? For example, could our system ask the analyst to respond to a prompt regarding the quality of the comments, which would subsequently guide adaptation of the software?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.
154	Are comments or approaches focused on the intelligence analyst's report history within the program scope?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.
155	Can information about the intelligence analyst, for example experience level, be requested from the analyst?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.
156	Does the analyst workflow allow for the capture of multiple interactions with the reasoning bot, understanding which corrections were accepted and which were declined?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.
157	Does the analyst workflow allow for any direct analyst feedback using natural language to the suggestions provided?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning. If the technical approach for TA3 can accept natural language feedback while meeting the requirements stated in II.A.8.b then it is permissible. Those requirements are "For TA3 Final Exams, the portion of Performer software facing analysts will be required to run within a word processing application. For planning purposes, Offerors may assume that this will be a Microsoft Word 365 add-in. These TA3 components will be permitted to access containerized TA1 and TA2 components running on the same testbed. They will not be permitted to access resources outside of the testbed."
158	If a tester uses the REASON tool repeatedly on the same analyst task, can the tool track it and make use of the information?	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.

159	Is it in scope for the analyst to filter commentary from the REASON application by indicating their priorities for comments	The only REASON evaluations with human users will be the TA3 Final Exam in each Phase. Offerors should not assume that there will be sufficient interaction with any single user to provide reinforcement learning.
160	Does IARPA desire a real-time interactive bot that assists analysts with reasoning, or only one which utilizes a word processing document as the interface?	As stated in the BAA (Amendment 1 of 5 APR 23) Section II.A.1 page 6: "REASON technology will fit into the existing intelligence analysts' workflow"
161	For the TA3 final exam with humans writing/revising reports, is it allowable for the performer system to present some number of comments when the writer asks for comments, and then subsequently present additional comments when the writer asks for more comments? That is, does the performer system have only one opportunity to present all comments at once to the writer, or can it present batches of comments upon writer request?	The Performer system for the TA3 Final Exam may present batches of comments upon request.
162	Can we think of this system as an iterative process, where comments are provided, amendments are made to the draft, and the process is repeated?	For the TA3 Final Exam and eventual operational use, yes, this is consistent with the Government's concept for REASON use.
163	Does the application have to enable the UX in the MSWord environment (e.g., plug-in)? Is an external application in scope that ingests the draft analytic report from Word, provides comments and allows the analyst to make modifications in an external application, and then exports an updated product back into the MSWord environment?	If the technical approach for TA3 meets the requirements stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b pages 16-17, then it is permissible. Those requirements are "For TA3 Final Exams, the portion of Performer software facing analysts will be required to run within a word processing application. For planning purposes, Offerors may assume that this will be a Microsoft Word 365 add-in. These TA3 components will be permitted to access containerized TA1 and TA2 components running on the same testbed. They will not be permitted to access resources outside of the testbed."
164	The proposal states: "REASON application deliverable must be compatible with analyst authoring applications, typically Microsoft Word." Does this set an expectation that it will be fully integrated with a software like Word and also be able to plugin elsewhere, or can it be separate but allow for easy movement between the two applications?	If the technical approach for TA3 meets the requirements stated in BAA (Amendment 1 of 5 APR 23) Section II.A.8.b pages 16-17, then it is permissible. Those requirements are "For TA3 Final Exams, the portion of Performer software facing analysts will be required to run within a word processing application. For planning purposes, Offerors may assume that this will be a Microsoft Word 365 add-in. These TA3 components will be permitted to access containerized TA1 and TA2 components running on the same testbed. They will not be permitted to access resources outside of the testbed."

165	<p>Each report that we receive for evaluation will contain significant context around which it is written. For instance, the language in a report concerning HUMINT will be different than that around TECHINT. The level of analytical rigor required may also vary significantly between different report types as domain knowledge and levels of certainty change. As a result, we can have a mismatch of report rigor required if a model is generated off of HUMINT reports, and then applied to TECHINT reports (HUMINT and TECHINT again as examples). We understand that performers will develop and test their systems against analytic reports provided by IARPA, and then T&E will test the systems against reports that the performers have not seen. Can we assume that in a single cycle, all reports (those that we receive and those that T&E uses) will belong to a single set (corpus) of report types? If so, in additional cycles of testing, does the system need to work for the prior sets equally as well as it had in prior tests? Can the system have different models for different report types?</p>	<p>Government-provided information is not intended to be the sole source of training data for Performers. Performer data requirements are given in BAA (Amendment 1 of 5 APR 23) Section II.A.5.b on page 11. Challenge Problems in any evaluation cycle will not necessarily be of any given type. Performer systems must field Challenge Problems of a wide variety.</p>
166	<p>Are there any specific expectations regarding the system's adaptability to new or updated data sources (the BAA notes that the corpus will be updated)? Are there expectations regarding the system's ability to handle different types of analytical products?</p>	<p>The corpus will be updated with new information of the same kind in order to keep it current with world affairs, in much the same way that a subscriber to a media outlet would receive new information periodically. Systems must be able to process English textual information of the kind found in the IC.</p>
167	<p>For the REASON's targeted users, what average level (number of years) of IC analysis experiences can we expect?</p>	<p>The Government will not provide that information.</p>
168	<p>Are there unclassified instructional materials used for training new analysts? If so, please provide an unclassified example.</p>	<p>The Government will not provide that information.</p>