

Questions & Answers – MATERIAL IARPA-BAA-16-11

Responses to Questions #9 to #38

Note: Questions posted verbatim and in order of time received.

Q9: Are academics working in Qatar eligible to apply to this bid as subcontractors?

A9: Per page 24 of the BAA, Foreign entities and/or individuals may participate as a subcontractor in collaboration with a U.S.-based prime offeror. Foreign participants must comply with any necessary Non-Disclosure Agreements, Security Regulations, Export Control Laws and other governing statutes applicable under the circumstances. Proposers are expected to ensure that the efforts of foreign participants do not either directly or indirectly compromise the laws of the United States, nor its security interests. As such, offerors should carefully consider the roles and responsibilities of foreign participants as they pursue teaming arrangements.

Q10: Is there an issue with Chinese Nationals working in Hong Kong participating in a proposal?

A10: Per page 24 of the BAA, Foreign entities and/or individuals may participate as a subcontractor in collaboration with a U.S.-based prime offeror. Foreign participants must comply with any necessary Non-Disclosure Agreements, Security Regulations, Export Control Laws and other governing statutes applicable under the circumstances. Proposers are expected to ensure that the efforts of foreign participants do not either directly or indirectly compromise the laws of the United States, nor its security interests. As such, offerors should carefully consider the roles and responsibilities of foreign participants as they pursue teaming arrangements.

Q11: May an academic collaborator who is also an intermittent professor appointee with NIST in the computational Mathematics division participate in a proposal or would the affiliation with NIST be considered a conflict of interest?

A11: Potential conflicts of interest shall be addressed separate from the Q&As. Notification of potential conflicts of interest must be included in your proposal with sufficient detail for the government to make a determination.

See section 3.A.1 Organization Conflicts of Interest (OCI) in the BAA.

"Offerors are strongly encouraged to read "Intelligence Advanced Research Projects Activity's (IARPA) Approach to Managing Organizational Conflicts of Interest (OCI)", found on IARPA's website at: <http://www.iarpa.gov/index.php/working-with-iarpa/iarpas-approach-to-oci>." page 4 of the IARPA approach documents states:

"It is incumbent on an offeror that is submitting a notification of a possible OCI to fully document any real or perceived potential OCI situation, and to make a clear and cogent argument explaining why it believes that no actual OCI exists or the Government should grant a waiver. At a minimum, this notification should: a. Fully disclose all facts relevant to the existence of the potential conflict of interest; b. Provide an OCI mitigation plan; and c. Identify any relevant

contract numbers under which the potentially conflicting activity is being or has been performed."

Q12. How will domains be defined during training/development and (the one new domain) during the evaluation period? Just a word? ("Government") A description? ("Government policies and actions, whether actual or proposed") A set of domain-relevant documents? (in English? in the target languages?)

A12: Domains will be identified upon each query release, along with a brief description of the domain. Each query will be associated with a domain, or high level subject category. Section 1.A.1 on page 9 provides possible domains for inclusion: Business and Commerce, Science and Technology, Health, Education, Environment, Entertainment, Sports, etc.

Q13. On slide 22 of the Proposers Day Briefing it is indicated that 5k of the program documents would be released at kickoff, an additional 5k (for a total of 10k) after the dev cycle concludes, and an additional 5k (for a total of 15k) after the CLE Eval but before the CLE+S Final Eval. The analogous figure in the BAA (figure 4) does not contain these numbers. Is the order of magnitude of the number of program documents per cycle anticipated to be the same as was indicated in the Proposers Day Briefing? If not, can you provide some guidance as to the order of magnitude of program documents per cycle?

A13: The actual number of documents released per language will not substantially differ from what was briefed at the Proposers' Day.

Q14: The BAA states: "Performers for MATERIAL will be expected to work with domains of interest and adapt their algorithms to domains without domain-specific training data." An example given in the BAA of a domain label is "government", a one-word label. In the absence of domain-specific data, how will the program specify what a domain label actually means, since virtually any label is ambiguous to some degree? For example, if "Health" were used as a domain label, it could be referring to anything from detailed scientific discussions of gene sequencing to political debates about Obamacare. Would it be possible for IARPA to provide a small number of positive and negative examples to ensure program-wide understanding of the meaning of the domain label?

A14: A brief description of each domain will be provided at each query release period. Section 1.A.1 on page 9 provides possible domains for inclusion: Business and Commerce, Science and Technology, Health, Education, Environment, Entertainment, Sports, etc.

Q15: on page 16 the BAA states that during the practice and surprise language development periods, NIST will provide a Development Server for performers to establish their method on the first epoch of documents released. Can you provide some detail on the development server? Will performers be allowed to take data off the server? Are performers expected to do all development on the server? Will the server also be used for evaluation?

A15: The Development Server provided by NIST is intended only for scoring AQWV during the practice and surprise development periods. There will be no data other than scores to take off this server. Performers will use their own hardware to develop their methods. To avoid confusion, the BAA will be amended with the Development Server renamed Scoring Server.

Q16: Are there any constraints on how fast we must compute and present a result? The BAA specifies that possible research approaches may address how to limit running time and memory (see page 10).

A16: Any time constraints on submission of evaluation results will be part of submission requirements distributed by NIST to performers at the appropriate times in each Program Period. It is important to note that each performing system will need to optimize integration of multiple human language technologies to ensure that evaluation results are produced within the specified time frame and without any human intervention.

Q17: Is the example in Figure 1, "polio vaccine; DOMAIN=government" representative of the type of query text we'll see (or might it be more like TREC-style long-form queries, or something else)?

A17: TREC long-form style queries will not be used in this program. As stated in Section 1.A on page 7 of the BAA, "MATERIAL queries, by design, include an English query string (QS) which expresses the topic to be retrieved, and a domain label which narrows down the set of relevant retrievable documents."

Q18: The BAA states that "The optimal summary will maximize the confidence of the relevance judgment and minimize the time required to perform the judgment." Does this mean that the evaluation is not just F-measure on relevance, but also takes into account analyst time?

A18: The AQWV metric will be used to assess performance during both CLIR and CLIR+S evaluations. As stated in Section 1.B.1 on page 11, IARPA may use additional metrics and analyses to gain a deeper understanding of factors that contribute to system performance.

Q19: The BAA states that "Participation by key personnel and significant contributors is expected to exceed (25%) of their time. A compelling explanation is required for any variation from this figure." What constitutes a "compelling explanation"? For example, would it be acceptable for some faculty, who have a 9 month academic salary, to put less than 25% of their time on the proposal?

A19: Per Section 5.A.2 of the BAA proposed work plans shall demonstrate that key personnel have sufficient time committed to the Program to accomplish their described Program roles. A compelling explanation must reflect this.

Q20: The BAA states: 6.B.13: "Software deliverables must be deployable to cloud platforms for testing and production use. Technical approaches should generally avoid the following: requiring high-performance, special-purpose, or excessive quantities of virtual hardware not readily available in the cloud;" Will GPUs be available? Will large memory machines be available?

A20: IARPA will not provide any hardware or software to performers to develop their solutions. At the end of each period of performance, IARPA will make a cloud-based platform available to allow performers to deploy their software and associated deliverables.

Q21: For the data provided will there be overlaps between ASR training data and the MT bitext data? Any expectations of how much?

A21: No overlap between these data sets is envisioned.

Q22: In the proposal format, the data management plan occurs twice: once in the technical volume (30 page limit) as section L. Detailed instructions are given about what should appear in L. The BAA also states that the data management plan should appear in the attachments and should be 2-3 pages in length. Should it appear both places or only in the attachment? Is there a difference in what should appear in the two different locations (if it should appear twice)?

A22: Per page 34 of the BAA, the Data Management Plan, described in Section 4.B.1.c.L shall be included as Attachment 8. These pages will not count towards the Volume 1 page limit. The data management template is provided as APPENDIX I, page 65 of the BAA.

Q23: The Data Management Plan is discussed in Section 4.B.1.c.L seemingly as part of the 30 pg. Technical Proposal – however, it is also identified as an Attachment that would not count against the technical proposal page limit. Is the Data Management Plan intended to be in both the main technical proposal and an attachment?

A23: Per page 34 of the BAA, the Data Management Plan, described in Section 4.B.1.c.L shall be included as Attachment 8. These pages will not count towards the Volume 1 page limit. The data management template is provided as APPENDIX I, page 65 of the BAA.

Q24: For one end-to-end system, what is the program required maximum turn-around time between the system receiving the queries and returning retrieved documents with their English summaries? Or is it true that there is not this requirement as long as the performer could submit results within the month of CLE and CLE+S evaluation period specified in Table 3 of the BAA?

A24: Per Section 1.B.2 of the BAA, page 12: NIST will provide the evaluation schedule, the evaluation procedure guidelines, submission requirements (e.g., rules of participation, formats for system output to enable scoring, and scoring methods), and evaluation packs to performers at the appropriate times in each Program Period. Performers will be required to complete all evaluations at their site without any human intervention and within the time frame specified by NIST, and to deliver evaluation output to NIST for scoring in the required format. NIST will score the outputs and provide results to performers and IARPA. It is important to note that each performing system will need to optimize integration of multiple human language technologies to ensure that evaluation results are produced within the specified timeframe and without any human intervention. Exceedingly long response times will be of less value to the IARPA mission.

Q25: In the BAA it states that summaries should only include *ideas* that occur in the document being summarized. We assume that synonyms of words that appear in the document are the same idea. What if the summary were to introduce language describing the document: "The document describes...". Would this be allowed? What if the summary were to state ideas the document does not contain "The document does not describe...". Would this be allowed?

A25: Synonyms and paraphrases are acceptable for presenting an idea. Performers are not prohibited from clarifying the summary content as long as they adhere to the 100-word limit stated in Section 1.A.2 on page 9 of the BAA.

Q26: For the cover letters (Attachments B & C), what is intended by "Technical Area"? Would "Human Language Technology" be an acceptable entry?

A26: Yes.

Q27: Under 4.B.2.b Section 2, should we assume a fiscal year of "October 1 – September 30" for a three year period?

A27: No, a fiscal year is 12 months starting from October 1 and ending September 30 of each year. The performance period is separate from the fiscal year period, and proposed costs must support the performance period.

Q28: Can participants license corpora from the Linguistic Data Consortium (LDC)?

A28: Per Section 1.A.5, pages 10-11 of the BAA, Performers are encouraged to harvest monolingual, bilingual or comparable corpora, wordlists, linguistic descriptions or audio from the web to develop their technologies, but with certain limitations. Performers may not use proprietary data that cannot be shared after the evaluation. Additionally, the MATERIAL Program will not fund the acquisition of new data collections beyond those provided. Please also note, data harvested for development will be shared with other performers after the end of the Program Period in which they were used. IARPA will provide any applicable LDC corpora that is available for government use.

Q29: Is it permissible to use parallel corpora in which both halves of the corpus are in English (rather than English and another language)?

A29: There are no restrictions as to the type of corpus used. Per Section 1.A.5, pages 10-11 of the BAA, Performers are encouraged to harvest monolingual, bilingual or comparable corpora, wordlists, linguistic descriptions or audio from the web to develop their technologies

Q30: We here at Johns Hopkins University are preparing a bid for the IARPA MATERIAL project. There will be employees of Johns Hopkins University involved in the project, who are appointed in the Human Language Technology Center of Excellence. I want to confirm that these employees eligible to be part of the project. To be clear - these are not employees associated with the Applied Physics Lab. Note, that this issue was already addressed in the IARPA BABEL project.

A30: "Offerors are strongly encouraged to read "Intelligence Advanced Research Projects Activity's (IARPA) Approach to Managing Organizational Conflicts of Interest (OCI)", found on IARPA's website at: <http://www.iarpa.gov/index.php/working-with-iarpa/iarpas-approach-to-oci>." page 4 of the IARPA approach documents states:

"It is incumbent on an offeror that is submitting a notification of a possible OCI to fully document any real or perceived potential OCI situation, and to make a clear and cogent argument explaining why it believes that no actual OCI exists or the Government should grant a waiver. At a minimum, this notification should: a. Fully disclose all facts relevant to the existence of the potential conflict of interest; b. Provide an OCI mitigation plan; and c. Identify any relevant contract numbers under which the potentially conflicting activity is being or has been performed." See also answer to question 11.

Q31: What annotations will be included in the analysis packs that will be released after the CLS Test? In particular, for each test query, will the packs include the corresponding set of relevant documents?

A31: Per 1.A.4, page 10: After each evaluation, an analysis pack will be provided for each language that will include translations and transcriptions of a proportion of the text and speech documents in the domains rolled out each period.

Q32: Will the program make available to the performers a subset of the development documents with relevance labels during the development period (especially in Phase I, before any *analysis pack* had been released)?

A32: Documents with relevance labels to a subset of the queries will not be provided during the development period.

Q33: AQWV is the acronym for *Average Query Weighted Value* on page 11 but in other places (e.g., page 12) it seems to stand for *Actual Query Weighted Value*. Do these two terms refer to the same metric?

A33: Per Section 1.B.2.1, page 13, Actual Query Weighted Value (AQWV) is the mean QWV (across queries) for the system when running with detection threshold θ . The terms do refer to the same metric. See Amendment 001.

Q34: The BBA states that summaries "must be rendered text in a static format". Is it allowed to render part of the summary in rich format, such as color or different font sizes? Is it allowed to include graphics, e.g., word cloud in bubbles?

A34: Rendered static images of words, such as multiple colors, fonts, sizes and alignments are allowed. Graphical elements such as lines, arrows, and bubbles are not. The entire summary may be delivered as a graphical image, such as a jpeg file.

Q35: During development, are we allowed to use English-speaking annotators to perform sample summary relevance judgments?

A35: Yes. Please also see Answer #38.

Q36: For one end-to-end system, what is the program expected maximum turn-around time limit between the system receiving the queries and returning retrieved documents with their English summaries (e.g., in xxx minutes, in xx hours)? Or is it true that there is not this requirement as long as the performer could submit results within the month of CLE and CLE+S evaluation period specified in Table 3 of the BAA?

A36: Any time constraints on submission of evaluation results will be part of submission requirements distributed by NIST to performers at the appropriate times in each Program Period. It is important to note that each performing system will need to optimize integration of multiple human language technologies to ensure that evaluation results are produced within the specified time frame and without any human intervention.

Q37: Does the program allow the built system to harvest web data during CLE and CLE+S evaluations?

A37: Web data may be harvested at any time as long as it is shared with other performers after the end of the Program Period in which the data were used.

Q38: Under Page 10 I.A.5, to clarify our understanding, is crowd-sourcing allowed to be used to inform summary development during the CLE and CLE+S evaluations?

A38: Performers are allowed to use crowd-sourcing to inform summary development at any time prior to the CLE evaluation of the corresponding practice/surprise language phase of each program period. The BAA will be amended to reflect this.