As of March 20, 2014

Q63. (a) Will you consider using Other Transaction Authority (845 agreement) for non-traditional performers? (b) If so, can it still be part of a BAA proposed team mixed with FAR performers?

A63. (a) Yes. (b) Yes.

Q62. (a) What goal is MUSE trying to solve that differs from DARPA'S Cyber Grand Challenge? (b) Is the program interested in finding and fixing vulnerabilities such as buffer overflow?

A62. (a) MUSE is focused on specification inference, automated repair (independent of issues related to security), and program synthesis. (b) Vulnerability detection and repair (e.g., buffer overflow) is within scope.

Q61. (a) Is the goal to infer non-functional properties or functional properties or both? (b) Should the inference try to infer the intent of pieces of code, using big data techniques?

A61. (a) Both. (b) Yes.

Q60. You mentioned that the corpus should be "high fidelity", but what does that mean? That is, "high fidelity" with respect to what?

A60. Code should be reasonably well-structured and debugged, but not necessarily verified, or known to be fully free of defects.

Q59. Will proposers that focus on TA2-4 areas need to develop techniques and tools that analyze multiple languages and multiple types of errors, or can they exclusively target some specific domain such as dynamic languages and some specific type of errors such as string manipulation errors?

A59. The corpus will be multi-lingual, but TA2-4 performers are not obligated to avail themselves of this property, although if they don't the proposal should clearly state the technical strengths of their strategy.

Q58. We intend to go with a single language family project for now. Is this of interest?

A58. Yes, but please refer to question and answer #59.

Q57. It seems that the TA1 performer will have a very short amount of time to produce challenge problems after the demo meeting at the end of the first 2 phases. Is it envisioned that the TA1 performer will be in communication with the other teams before those meetings?

A57. Yes.

Q56. (a) How much teaming do you expect to coordinate after proposal submission vs. teaming before proposal submission? (b) Could sections of a proposal be selected for award (e.g., if an organization proposes to TA2 and TA4, might the government award work for TA2 but require teaming on TA4)?

A56. (a) Level of teaming will depend on the composition and merits of the proposals, and determination of synergies. (b) Yes.

Q55. We are potentially interested in applying MUSE findings to software for a novel application (i.e. robotics). Is this a valuable task to DARPA? We can apply findings to ongoing DARPA and DoD work and assess results.

A55. This may be valuable (in the abstract), but be aware that I2O does not have an active interest in robotics research.

Q54. Is there an example or boilerplate ACA available?

A54. No, but a boilerplate ACA clause is attached to this FAQ. Be aware that the ACA is only relevant after favorable selection.

Q53. Will all the source code work (modify oss, build new, etc.) for MUSE be required to be released back out to the open source software community?

A53. The infrastructure will be made available to the public after program completion.

Q52. Will all work occur in an open manner (i.e., occur on github so that all performers can view and comment)?

A52. Performers will have access to all materials via private wiki/cloud infrastructure, but this will not be available to the general public during the lifetime of the program.

Q51. How can the TA5 performer make sure that new analyses can be integrated easily if the TA4 performer is controlling the architecture? Is there some agreement that needs to be reached between them?

A51. The TA4 performer defines the software architecture of their AAS, not the overall system architecture of the program. In particular, they do not control the organization, management, and maintainability of the corpus and system resources.

Q50. Can you clarify the categorization of the corpus, especially the semantics and granularity? For example, the slide had strings as a category. Clearly that is a large space and string manipulation may (a) be part of diverse systems and (b) be built on a common core such as the c string library. How do these impact the categorization of the corpus?

A50. The TA1 performer will be responsible for defining semantically sensible ontologies that will facilitate TA2 analysis.

Q49. (a) How will you ensure that the artifacts needed by a TA4 proposer be generated by a TA2 performer and the right TA2 team will be paired with the TA4 team? (b) Is the TA4 team better off including artifact generation as a component (i.e., combined TA2 and TA4 proposal), as a minor component, or an option?

A49. (a) Our determination will be based on realizing best synergies among the teams based on the technical material presented in the proposals. (b) Proposals should be written exclusively on technical merit, and not on presumed advantages of teaming combinations, if such combinations weaken the overall strength of the proposal.

Q48. Does the bibliography count against page count for the technical section, or can it be put in Appendix B?

A48. The bibliography does not count against the page count for the technical volume if it is put in Appendix B. See Page 35 of the BAA.

Q47. (a) What will the award look like in terms of duration and ceiling? (b) Will it be a single MAC for all awardees?

A47. (a) The program is 52 months. (b) No.

Q46. (a) What is the timing of the ACA (associate contractor agreement)? (b) Does this have to be in place when the proposal is submitted, or after the performers are selected?

A46. (a) The ACA must be in place by program kickoff. (b) No.

Q45. Can proposals include DIST-C / ITAR data?

A45. Proposals may if that supports salient technical arguments.

Q44. Do you have any concrete use-cases in mind for how people might use the end system for application construction?

A44. This will be developed as part of the challenge problems provided by AAS teams and the TA1 performer.

Q43. Do you have any concrete examples in mind of complex inferences that might be derived from facts in the database?

A43. Inferences that enable learning of models are one such example described in the BAA.

Q42. Will the proposal submission mechanism allow submitting early and then submitting updates (until the deadline)?

A42. Yes.

Q41. How much human interaction is allowed (does it have to be fully automated or human guided)?

A41. There is no explicit restriction or bounds defined in the BAA on this issue.

Q40. How is it envisioned that TA3 and TA4 interact with regard to the specification language?

A40. Proposals will need to address this question explicitly.

Q39. How many contracts do you anticipate awarding in each TAs 2-4 and in which phases?

A39. This depends on the quality of the proposals submitted.

Q38. (a) Will past performance be evaluated? (b) Does it need to be formally submitted?

A38. (a) No. (b) No.

Q37. Are there requirements or quotas for the inclusion of small businesses?

A37. No.

Q36. Is each benchmark a single application or a domain such as sorting?

A36. Most likely a combination of both, but the exact structure will be determined by the TA1 performer.

Q35. Do applications in the corpus have to match the challenge problem precisely or can they implement parts of a challenge problem?

A35. Code in the corpus will not match the challenge problem precisely (if it did, then we'd be doing search, not synthesis or repair).

Q34. The benchmarks and challenge problems will presumably be written in a number of programming languages. (a) Are TA2 performers expected to provide analyzers for all of these languages? (b) Is it required to analyze all benchmarks and CPs?

A34. (a) No. (b) No.

Q33. What is the expected size of a TA2, TA3, or TA4 team in terms of budget, number of co-PIs, institutions, etc., particularly from an academic institution?

A33. This depends on the technical scope and breadth of the proposal.

Q32. Can you clarify the rules about PIs participating in more than one proposal? My understanding is that they cannot participate in more than one proposal in a single technical area, is that correct? Not even as a subcontractor? But a PI may participate in several proposals as long as each proposal is for a different technical area and the technical areas are not in conflict.

A32. Correct, a PI can participate in several proposals as long as each proposal addresses a different technical area.

Q31. Suppose two entities want to submit a collaborative proposal. Does this have to be done as a subcontract, or can each entity submit their own proposal and indicate in the proposal that it is a collaboration?

A31. The choice depends on whatever arrangement would work best for the organizations involved. However, each proposal should be able to stand on its own merits.

Q30. Should universities use grants.gov? What is the type of agreement for universities?

A30. If you are submitting a cooperative agreement, use Grants.gov, otherwise use the DARPA submission portal.

Q29. To what extent does scalability matter to the overall program, especially to TA2 and TA4? There may be practical limits to the amount of time that may be taken for an analysis in order for data mining to be effective.

A29. The program is very much focused on scalability. The practical limits raised in the question should form part of the proposal narrative if such an issue figures prominently in the technical solution proposed.

Q28. In TA1, it states that "the corpus may also include annotated and verified programs". Are these meant to be "ground truth", and annotated accordingly?

A28. Yes.

Q27. How are the enclaves discovered by TA4 different or related to the ontology categorization provided by TA1?

A27. The ontologies defined by TA1 are not based on deep semantic analysis or mining, but organized around either superficial examination of the program, or on specific domain knowledge.

Q26. Can the TA3 database be pre-populated with domain-specific or other knowledge (e.g., programming language models, abstract operation sequences)?

A26. Yes.

Q25. (a) Can TA2 mine artifacts it generates for a program, including any additional information available about it to make new inferences about that program? (b) Can TA2 also query the graph database, not just populate it?

A25. (a) Yes. (b) Yes.

Q24. Can TA4 (or TA2?) detect certain program bugs based solely on program analysis -- without requiring any data mining -- when such detection is possible? Some bugs may depend solely on how that program is implemented, and may not be amenable to detection using mining techniques.

A24. Yes.

Q23. Is it a requirement that every repair be a provably correct repair?

A23. No.

Q22. Is it within the scope of the program for TA4 to allow human feedback (e.g. labeling) with regard to discovered patterns?

A22. Yes.

Q21. (a) Can TA2 use COTS tools for analyzing programs? (b) If yes, should licensing costs for such tools be included in the cost proposal?

A21. (a) Yes. (b) Yes.

Q20. (a) Can TA2 propose techniques for a subset of programming languages? (b) Do both source and binary analysis need to be addressed by each TA2 proposal?

A20. (a) Yes. (b) No, although be aware that the corpus will likely include both.

Q19. (a) Will corpus program be self-contained (i.e., have everything needed to run them)? (b) Will programs need to be run manually, or will there be automated test scripts?

A19. (a) Yes. (b) The TA1 performer is expected to provide test scripts.

Q18. If we are submitting a proposal for more than one of TA2-4, should the cost proposal have separate costs for each TA?

A18. Yes, but include in the cost volume an explanation of potential cost efficiencies realized by awarding multiple TAs.

Q17. If submitting proposal for more than one of TA2-4, should the technical plan section have separate subsections for each TA?

A17. This depends on how you believe your technical ideas can be best conveyed.

Q16. In your presentation you mention program quality, program synthesis, program attributes and program repair. Is your use of the word "program", limited to the source code, or do you consider this to be a broader definition, which can include other program artifacts such as test cases, documentation and the architecture? For example, do you envision test case synthesis, test case repair, and/or quality? We could ask the same thing about other artifacts.

A16. To limit the scope of the program (which is already quite broad), we are not considering artifacts such as test cases, architectures, etc. as targets of mining and repair.

Q15. It seems that there's a large (100%) focus on formal specifications -- is there any interest in informal specifications, e.g., those written in natural language that one might find in comments, spec sheets, bug reports/fixes, etc...?

A15. The BAA does admit non-formal methods (such as documentation analysis) as acceptable techniques and sources for new inferences.

Q14. Can you give a small, concrete example of a few fact-nodes that might appear in the database, a few edges that might connect them, and an interesting analysis that might use these facts and relationships to answer a specific question that a programmer might have?

A14. Facts could be types and pre/post conditions. Edges connect facts discovered by different analyses for the same program or method. Abductive inference is an example analysis that could benefit from this structure.

Q13. Could two proposals from a university (with different PIs) both be selected if one was TA1 and the other was TA2-4? Could a university be selected as a prime on a TA2-4 proposal and a subcontractor on a TA1 or TA5 proposal?

A13. No. To ensure independence and prevent conflict of interest, proposers selected for TA1 or TA5 will not be selected for award as performers (prime or subcontractors) on TA2, TA3, or TA4.

Q12. Are TA1 performers allowed to utilize "TA1 relevant" work products (software corpus, ontologies, etc.) created prior to award by organizations that are competing for TA2-4, or will that result in a conflict of interest?

A12. This is discouraged. If such products are included, then justification that these products do not incur a conflict-of-interest or provide undue advantage to the TA2-4 performers who developed the product must be provided.

Q11. What assumptions can TA2 proposers make about the scope of benchmarks, e.g., in terms of the programming languages as analysis tools are generally restricted to specific programming languages?

A11. They may assume the corpus will include a broad class of well-studied and well-known languages and systems.

Q10. How can a TA2 performer estimate the cost if the benchmarks will be identified after the program kickoff?

A10. Costs are to be determined only on the basis of the technical approach taken, which should not be tailored to specificities of benchmarks or challenge problems.

Q9. Is it preferable to select a proposal that combines TA2-TA4 or will proposals that focus on a single area have an equivalent chance for getting funded?

A9. There is no predetermined bias on favoring integrated vs standalone proposals.

****

As of March 12, 2014

Q8. I would like to submit a proposal that encompasses different technical areas (TAs). Can this be done?

A8. Yes. Proposers may address any of the five technical areas, which are discussed in the BAA. To ensure independence and prevent conflict of interest, proposers selected for TA1 or TA5 will not be selected for award as performers (prime or subcontractors) on TA2, TA3, or TA4. The decision as to which technical area(s), if any, to consider for award is at the discretion of the Government.

While proposers may submit a proposal on any of the five technical areas, proposers may also address more than one TA in their proposal. Please note that due to the nature of the work under each TA, proposers may combine the following technical areas in a single proposal: TA1 & TA5; TA2 & TA3; TA2 & TA4; TA3 & TA4; and TA2, TA3 & TA4. No other combination of technical areas should be attempted in a single proposal.

Q7. If I submit a proposal that encompasses different TAs, as described above, will I be allowed to use more pages in Volume 1 of the proposal?

A7: Yes. Should a proposer submit a single proposal that encompasses two TAs, the maximum page count of Volume 1 (Technical and Management Proposal) is increased to 33 total pages, including all figures, tables and charts, but not including the cover sheet, table of contents or appendices. If a proposer submits a proposal for TA2, TA3 & TA4 (three technical areas), the maximum page count of Volume 1 is increased to 40 total pages, including all figures, tables and charts, but not including the cover sheet, table of contents or appendices.

Q6. Is each college/department within a university allowed to submit a proposal as a prime towards Technical Areas 1-5?

A6. Separate colleges, departments, and research groups at universities represent separate entities and, thus, are allowed to submit their own individual proposal towards TA1-5.

Q5. Confirm that TA2/TA3/TA4 will have to wait 6-months post kick-off meeting prior to getting the TA1 benchmarks for system development leaving 8-months to accomplish all phase 1 work as deliverables are due one month prior to Demonstration Workshop?

A5. Yes, that is correct.

Q4. Will an organization be excluded if intellectual property rights are not in alignment with open source regimes?

A4. No, but per the BAA, if proposers desire to use proprietary software or technical data or both as the basis of their proposed approach, in whole or in part, they should: 1) clearly identify such software/data and its proposed particular use(s); 2) explain how the Government will be able to reach its program goals (including transition) within the proprietary model offered; and 3) provide possible nonproprietary alternatives in any area that might present transition difficulties or increased risk or cost to the Government under the proposed proprietary solution.

Q3. What is the process that the Government will use to establish design teams?

A3. Fully integrated proposals spanning TA2-4 will be determined exclusively on their merits alone. For individual proposals in TA2-4, the proposers will be grouped into AAS teams based on the Government's determination of best synergies.

Q2. Define the baseline of what the Government considers the existing state of practice?

A2. Generally available publications and tools.

Q1. Please provide a total amount of funding available and what the Government believes is the distribution across the Technical Areas?

A1. The overall program budget is roughly $45M. The rough distributions across the technical areas are as follows:

- 10-15% for TA1;
- 75-85% for TA2-4; and
- 5-7% for TA5

ATTACHMENT

Clause Example - Associate Contractor Agreements

(a) It is recognized that success of the (Program name) research effort depends in part upon the open exchange of information between the various Associate Contractors involved in the effort. This clause is intended to insure that there will be appropriate coordination and integration of work by the Associate Contractors to achieve complete compatibility and to prevent unnecessary duplication of effort. By executing this contract, the Contractor assumes the responsibilities of an Associate Contractor. For the purpose of this clause, the term Contractor includes subsidiaries, affiliates, and organizations under the control of the contractor (e.g. subcontractors).

(b) Work under this contract may involve access to proprietary or confidential data from an Associate Contractor. To the extent that such data is received by the Contractor from any Associate Contractor for the performance of this contract, the Contractor hereby agrees that any proprietary information received shall remain the property of the Associate Contractor and shall be used solely for the purpose of the (Program name) research effort. Only that information which is received from another contractor in writing and which is clearly identified as proprietary or confidential shall be protected in accordance with this provision. The obligation to retain such information in confidence will be satisfied if the Contractor receiving such information utilizes the same controls as it employs to avoid disclosure, publication, or dissemination of its own proprietary information. The receiving Contractor agrees to hold such information in confidence as provided herein so long as such information is of a proprietary/confidential or limited rights nature.

(c) The Contractor hereby agrees to closely cooperate as an Associate Contractor with the other Associate Contractors on this research effort. This involves as a minimum:

> (1) maintenance of a close liaison and working relationship;
> (2) maintenance of a free and open information network with all Government-identified associate Contractors;
> (3) delineation of detailed interface responsibilities;
> (4) entering into a written agreement with the other Associate Contractors setting forth the substance and procedures relating to the foregoing, and promptly providing the Procuring Contracting Officer with a copy of same; and
> (5) receipt of proprietary information from the Associate Contractor and transmittal of Contractor proprietary information to the Associate Contractors subject to any applicable proprietary information exchange agreements between associate contractors when, in either case, those actions are necessary for the performance of either.

(d) In the event that the Contractor and the Associate Contractor are unable to agree upon any such interface matter of substance, or if the technical data identified is not provided as

scheduled, the Contractor shall promptly notify the DARPA (Program Name) Program Manager. The Government will determine the appropriate corrective action and will issue guidance to the affected Contractor.

(e) The Contractor agrees to insert in all subcontracts hereunder which require access to proprietary information belonging to the Associate Contractor, a provision which shall conform substantially to the language of this clause, including this paragraph (e).

(f) Associate Contractors for this (program name) research effort include:

Technical Area	Contractor Name and Point of Contact (POC)
TA1	
TA2	
TA3	
TA4	
TA5	

(end of clause)