

DARPA-BAA-16-53
Explainable Artificial Intelligence (XAI)
Frequently Asked Questions

As of August 26, 2016

Q51: Is the user a novice or domain expert? How much domain knowledge and experience would the end user have?

A51: The user is a domain expert, but not a machine learning or AI expert. Variations in user domain expertise are not critical to the program. During Phase 2 we will select subjects who possess average expertise in the problem domain. It is important to provide explanations that enable interaction and drill-down, but the explanation does not need to be tailored to the user's expertise.

Q50: Is it important to explain uncertainty?

A50: Yes. It is expected that it will be important for the system to convey its uncertainties and confidence level to give the user a better understanding of the system's strengths and weaknesses.

Q49: Creating an explanation can require an understanding of the particular needs/mental state/knowledge of a user. Where is this issue part of the XAI program relative to TA1 and TA2?

A49: Both. In general, both TA1 and TA2 require expertise in psychology and cognitive science. TA1 requires application of this knowledge to design and develop an effective user interface. TA2 requires more theoretical knowledge to construct a computational theory. However, the program will not emphasize tailoring explanations to user variations in experience or knowledge. It is important to provide an explanation that enables interaction and drill-down, but the explanation does not need to be tailored to the user's expertise.

Q48: Are "one shot" explanations ok? Or is a capability to have a dialogue about the explanations required?

A48: Sure, if you can provide the perfect explanation in one turn. We expect that some interaction and drill down will be needed to provide an effective explanation, but it is not mandatory.

Q47: Is the “goodness” of an explanation dependent on what the human knows and accepts to be true? Will there be different levels of expertise in the evaluations?

A47: No. Variation in the user’s domain knowledge is not critical to the program. We will not have different levels of user expertise in the evaluations. Moreover, the explanation does not need to persuade the user that the explanation is correct, just provide an understandable rationale for the machine’s decision.

Q46: Is it necessary for the explainable visualization to be fine grained, high level or somewhere in between?

A46: It depends on the situation. We expect it will be important to provide an explanation that enables interaction and drill-down as needed, including adjusting the coarse/fine graininess of the visualization.

Q45: Is research on how trust is created between machines and people relevant to the BAA? It’s a bit different than the psychology of explanation. Does it fit in that area?

A45: No. Research specifically on how explanations create trust would be in scope, but the general problem of human-machine trust is out of scope.

Q44: Can a team focus on explanation interfaces on existing ML solutions (rather than brand new ML techniques)?

A44: Yes.

Q43: Which is weighted more, the explanation model or the HCI component?

A43: They’re weighted equally. We are looking for the effectiveness of the overall system. We will look at the complexity of the problems going in, and the quality of the explanation coming out. Innovation is important for both components.

Q42: Can the modified ML techniques be based on SVM?

A42: Yes.

Q41: Is there any interest to bring perspectives from control theory and distributed control and optimization into reinforcement learning and inverse RL?

A41: Yes. From the BAA: “Because the decision machinery is likely to include additional planning, decision, or control modules, DARPA anticipates that these explanations may cover those reasoning modules, as well as decision policies learned through reinforcement learning.”

Q40: Are swarms or ensembles of autonomous entities permissible/in scope of the program or do you want a single entity that will be explaining itself?

A40: Swarms are within the domain of the program. A swarm acting as a single agent is acceptable in the program.

Q39: What is the amount of background knowledge and pre-processing that a system can do prior to an evaluation?

A39: A reasonable amount. We understand that some background knowledge may be necessary, but we want to avoid a manually intensive knowledge engineering effort.

Q38: Can you elaborate on the evaluation protocol? Is it “over the fence” to the evaluator or is the TA1 researcher there to assist and adjust?

A38: It is NOT an “over the fence” evaluation. The exact protocol and interfaces for the evaluation will be defined by the government evaluator during Phase 1. But, we expect that the evaluation will be accomplished through a web interface. The performer will provide access to the user interface via a web page that the evaluator can access. The performer will provide a data input API as defined by the evaluator. The system will be running in the performer’s computing environment during the evaluation. Performers will NOT deliver software “over the fence” for the evaluator to run in their environment.

Q37: Do the challenge problems proposed by developers in Phase 1 need to be as described in the BAA (i.e., multimedia analytics and UAV autonomy)?

A37: Proposers should suggest creative and compelling test problems that would be the most productive drivers of XAI research and development. The program seeks test problems that are sufficiently general and compelling to be useful for multiple XAI approaches in order to avoid having a unique, tailored problem for each research project. There are several questions regarding the exact nature of the data set or autonomous system that should be proposed for Phase 1 and how closely these test problems should match the problems described in the BAA. We have a preference for problems as described in the BAA, but that is not a hard requirement. These are the desired properties for the test problems proposed by the TA1 developers, in priority order:

1. Propose creative and compelling test problems that would be the most productive drivers of XAI research and development.
2. Propose test problems in one or both of the two general categories described in the BAA (data analytics or autonomy).
3. Propose Phase 1 test problems, data sets or simulation environments that closely match the desired characteristics of the two problem areas (a data analytics problem that involves heterogeneous, multimedia data and an autonomy problem that involved a simulated, autonomous UAV).

Q36: Do data analytics data sets need image/video data or will heterogeneous data with no video component be considered?

A36: There is a preference for heterogeneous multimedia data, especially by the time we reach the end of Phase 2, but it is not an absolute requirement.

Q35: Can speech be a standalone test modality, or should it be combined with text/video for multimodal event detection?

A35: Like Multimedia Event Detection – it is preferred to have multimedia data. However you can propose, if compelling, a standalone speech test modality for Phase 1.

Q34: For Phase 1, could your test problem address speech and text separately, and then be combined in Phase 2 for a multimedia problem.

A34: Yes.

Q33: Can proposers address the autonomy challenge with systems other than UASs (autonomous cars, non-military systems, etc.)?

A33: Yes.

Q32: Is there an option for Phase 2 to use classified datasets even if the software remains unclassified?

A32: No. The program is unclassified.

Q31: Do the challenge problems proposed by developers in Phase 1 also have to be defense-related?

A31: No.

Q30: Has DARPA identified possible Department of Defense transition partners, and if so, can you share who they are?

A30: Informal discussions are ongoing. A number of transition agencies are interested in the technology and will follow progress of the program. As the technology develops and matures, we expect to engage more actively with these transition partners.

Q29: Who conducts annual evaluations? Performers or T&E?

A29: In Phase 1, the Performers will conduct the evaluations; and in Phase 2, the government evaluators will primarily conduct the evaluations with assistance from the performers.

Q28: Will developing data sets or simulations for the TA1 evaluation be in scope?

A28: Yes, but don't spend too much money.

Q27: Can classified data sets be used in Phase 1 challenge problems?

A27: No. The program is unclassified.

Q26: Can social media data sets be used in Phase 1 challenge problems?

A26: Yes, assuming you are careful to not invade anyone's privacy.

Q25: Are there expectations for platform compatibility and processing times?

A25: No. During the learning phase you run it in whichever environment you choose. During the evaluation, the learner will be plugged into an evaluation platform. The government has access to high performance computing – if you would need access to it you should state so in your proposal. Processing times will not be part of the evaluation criteria.

Q24: Do you have a preference for the size of the training set/platform (as social media data sets are quite large)?

A24: A specific size of the training set/platform has not been determined, however, data sets like those in Kaggle or NIST challenge problems have been considered.

Q23: Will applications be looking for datasets on the order of petabytes?

A23: No. In the real world, maybe. But we are not looking to scale it up to petabytes right now. Just have the set large enough so that the machine learning challenge is adequate.

Q22: Who is the government evaluator?

A22: DARPA has an evaluator candidate in mind, but nothing has been finalized at this time.

Q21: In terms of evaluation for Phase 1 (where we choose the challenge problem), would the government evaluator dictate the evaluation policy, or would these be dictated by the performer?

A21: Phase 1 evaluation policies will be dictated by the performer.

Q20: For evaluation, will we have to implement our own evaluation framework (following an agreed upon protocol), or would this be solely implemented by the evaluator?

A20: For Phase 1, it would be your own evaluation framework. Hopefully, roughly similar to the evaluation framework for Phase 2 for compatibility purposes.

Q19: Is it ok for a performer to stay within only one of the general problem categories?

A19: Yes.

Q18: Is it ok for one team to focus on one challenge problem for the duration of the program.

A18: Yes that's how the program proposals will work. You should declare which challenge problem you are proposing to, and you will work on that problem for the duration of the program.

Q17: How extensive do you expect/want human in the loop experimentation to be included in the proposal?

A17: It is important to have human in the loop evaluations – but with a reasonable progression. It would be reasonable to begin with more frequent, less formal, pilot evaluations and move to larger more formal evaluations later in the program.

Q16: Is sample size important or is just a small pilot test sufficient?

A16: Small pilots are fine in the beginning, but larger sample size will be important later on in the program.

Q15: Are you looking for integration teams? What size?

A15: We are looking for proposers, as described in the BAA under TA1, to develop prototype explainable learning systems, which would include both a machine learning technique that produces an explainable model and an explanation interface. We are NOT looking for a broader integration team to pull together all of the TA1 technological solutions. At proposers' day, we presented a slide regarding team size and composition:

TA1: Explainable Learners

- Each team consists of a machine learning and a HCI PI/group
- Teams may represent one institution or a partnership
- Teams may represent any combination of university and industry researchers
- Multiple teams (approximately 8-12 teams) expected
- Team size ~ \$800K-\$2M per year

TA2: Psychological Model of Explanation

- This work is primarily theoretical (including the development of a computational model of the theory)
- Primarily university teams are expected (but not mandated)
- One team expected

Q14: Are there going to be large integration teams who combine all directions?

A14: No.

Q13: Can we compose a team as we feel?

A13: Yes. The term “team” can be a single organization or be composed of several different organizations.

Q12: Is there a preference for universities, industry labs, or both?

A12: There is no preference. It is up to the proposer to determine the best approach that leads to the most innovative and effective research and development effort to address the goals of XAI.

Q11: Can an entity be part of several proposals?

A11: Yes.

Q10: Is there a limit on the number of proposals per institution?

A10: No.

Q9: Will PI meetings be in DC or throughout the US?

A9: As described in the BAA, we will start in Arlington, Virginia, and tentatively anticipate alternate meeting locations between East Coast and West Coast after that. This information is for budget planning purposes only.

Q8: What is the target date for feedback on abstracts?

A8: Before October 1, 2016 (but hopefully much earlier).

Q7: Will DARPA recommend teaming partners after abstracts are submitted?

A7: No. It is up to the proposer to determine the best team that leads to the most innovative and effective research and development effort to address the goals of XAI.

Q6: Would you like to have teams established when abstracts are submitted or when proposals are submitted?

A6: You can propose an abstract that only has less than a full team. It may affect the type of feedback received. Feedback may focus on the facets of the program that you did address in the abstract. As a result, it may be beneficial to have an abstract that includes a whole team because you will receive more comprehensive feedback.

Q5: Will you post the Proposers Day slides?

A5: Yes. The Proposers Day slides are posted to the DARPA website, and a recording of Proposers Day will be posted in approximately two weeks after all material has been cleared for public release. All of this information is/will be posted on the DAPRA Opportunities web page: <http://www.darpa.mil/work-with-us/opportunities>.

Q4: Who are the contracting agencies?

A4: Several Department of Defense agencies (SPAWAR, AFRL, and possibly others).

Q3: Do we have to justify doing all the work in-house (no subcontractors) under all circumstances, or was that just about a particular type of proposal or scenario?

A3: Submission of a subcontracting plan or a justification for no subcontracting possibilities applies to any non-small business proposing a FAR-based contract that would exceed \$700,000. (See: FAR 19.702)

Q2: Will you issue grants?

A2: Grants will not be awarded under this solicitation. Procurement Contracts, Cooperative Agreements, or Other Transactions (OTs) may be awarded.

Q1: What are the procedures/limitations on publications for non-academic institutions?

A1: See under BAA Section II.B "Fundamental Research" the publication restriction language that would be incorporated into any resultant non-fundamental research procurement contract or OT.