Security and Privacy Assurance Research (SPAR) BAA Questions

#	Question	Answer	Date Posted
001	Are those who provided technical consultation on another IARPA program eligible to submit proposals to the SPAR program?	An organization with an employee serving in a “technical consultant” role on a specific program may be eligible to perform on other IARPA programs subject to a determination that its employee’s performance as a technical consultant did not give it access to government or proprietary information that would give it an unfair competitive advantage with respect to the IARPA program in which it seeks to compete. This determination must be sought through an Organizational Conflict of Interest (OCI) waiver request, as described in section 3.A.1 of the BAA.	2/01/11
002	Is information available about current participants and interested parties in the SPAR BAA. Was there an industry day? Is there any additional information available beyond the BAA, the Appendices, and the Interested Vendors List?	There are no current participants as this is a new BAA. Interested parties may post entries on the FBO Web site. An Industry Day was not held. All available information is posted to the FedBizOpps Web site at: https://www.fbo.gov/index	2/01/11
003	When is the last date that IARPA will accept questions about the BAA?	Section 7 of IARPA-BAA-11-01 notes that questions will be accepted until 3 February 2011, which is 15 days before the proposal due date of 18 February 2011 for the initial round of selections. After the due date for the initial round of selections, questions will again be accepted from 19 February 2011 until 13 December 2011, 15 days before 28 December 2011, which is the last day the BAA itself will be open. As noted in paragraph 4.A.1, the BAA will remain open for one year from release, and all compliant proposals received while the BAA remains open will be evaluated.	2/04/11

#	Question	Answer	Date Posted
004	What alternatives are there for “type of procurement contract” (section 1, coversheet item 10)?	Procurement contracts may include cost reimbursement, fixed price, and time and materials contracts. It is anticipated that a cost-reimbursement type contract is appropriate to address the uncertainties in fundamental research, but this determination will be addressed by the contracting officer during contract negotiations.	2/04/11
005	What level of detail is required in the total cost breakdown? Specifically, is it sufficient to include individual salaries of employees and associated federal and state taxes or do we have to itemize health/disability insurance/workers comp etc etc. (4.B.2, section 2, coversheet item 1)?	The Federal Acquisition Regulation (FAR) requires that you provide sufficient information to allow the contracting officer to determine that the awarded contract price is fair and reasonable. You must identify all costs associated with performance of the contract, including all direct and indirect charges. The Government must be able to verify the charges. The level of detail required will depend somewhat upon the type of contract and the total dollar value. However, since you will have to identify all such costs to prepare an executable proposal, the Government simply requires that you provide that information for review. As your costs would include health, disability insurance and all other indirect costs, your proposal would not be complete without including and itemizing those costs.	2/04/11
006	What “cost or pricing data” is the second to last paragraph in 4.B.2 referring to? Is it referring to the amount requested for the IT and equipment purchases? Or for the entire contract?	The definition for "cost or pricing data" is provided in FAR 2.101, which has an extensive description of the term. Any amounts you propose for IT or equipment purchase would be included in the cost or pricing data, along with direct labor, indirect costs, materials, and any other element of cost included.	2/04/11
007	Regarding data set usage: I assume if the data sets to be used are synthetically generated, there are no concerns regarding the discussion in section 3.D.	No. The requirements sets forth in Section 4.B.1/Section 3.D (page 46) of the BAA are relevant to both synthetically generated and “real” data sets.	2/04/11

#	Question	Answer	Date Posted
008	In 1.A.3.3.3: “Efficiency will be evaluated relative to a comparison system such as Apache Hadoop(TM) or a similar system to be selected by the Government.” Several of the operations implied by TA3.3-R1 and TA3.3-R7 are not supported by Hadoop or its standard libraries, so it is hard to imagine how a comparison with Hadoop would work for all of the criteria. Would the efficiency requirement TA3.3-R8 be measured only for the operations that Hadoop supports, or all operations? Is is possible to provide a list of candidate “similar systems”, or an acceptable API that the system must provide?	Apache Hadoop (TM) supports the operations required by TA3.3-R1 and TA3.3-R8, namely, retrieval, insertion, removal, and modification of stored items. For example, the Hadoop Distributed File System (HDFS) and the Apache HBase project both support all of these operations. The efficiency requirement TA3.3-R8 will be measured for each of the named operations. TA3.3-R7 addresses only retrieval operations and requires support for complex queries chosen from those complexity features listed in Table 4 of the BAA. Innovative approaches are sought that protect security and privacy in the context of practical range of data retrieval complexities. Apache Hadoop is mentioned as an example of a data storage capability that scales to very large data sets, and is widely available. There is no list of similar systems or APIs available. Note that the specific features of any named example comparison system should not be viewed as requirements or acceptable limitations on the proposed research prototypes.	2/04/11
009	In TA3.3-R8, the metric is the “mean elapsed time”. Depending on the underlying network, if the latency of each operation is limited by the latency and/or bandwidth of the network, then the requirement that the TA3.3 system perform operations in no more than twice the time required by the evaluation system has important implications about the number of message exchanges in the TA3.3 protocol relative to the protocol used by the evaluation system. Will the evaluation testbed have a low-latency, high-bandwidth network between the server and the client (1 GbE or faster) or does this imply that the network will be the bottleneck?	The Test and Evaluation environment will be designed to minimize the sensitivity of measurements of the prototype and comparison systems to uncontrolled sources of noise. The exact experimental design is yet to be determined by the Government.	2/04/11

#	Question	Answer	Date Posted
010	For the pub/sub capability in TA 3.1, is it assumed that new subscribers to a topic will only be able to receive messages posted to a topic after they subscribe to it, or should new subscribers be able to receive all messages ever posted to a topic? Should we make any inference about when the server may delete items or whether a client that subscribes now to some interest should receive messages previously sent to that interest? In TA 3.1, suppose several clients have submitted a number of items for some interest over a period of time. Suppose that another client then subscribes to that item. Is the server required to publish the previously submitted items to the new subscriber? In TA3.1, is there any functional requirement that would necessitate a server retaining an item after it has been published to all of the current subscribers to the interests that the item relates to?	TA 3.1 only requires that an item be available to Clients with matching subscriptions that are active at the moment the item is received by the Server. Conceptually, the items may be assumed to pass through Server processing in a continuous stream with no further access to items after their processing. Although this is not explicitly stated, it is implied by requirements TA3.1-R4 and TA3.1 R5.	2/04/11
011	The proposal does not state monetary amounts that are associated with each task. Can you please give some indication of what would be considered reasonable for each task?	The anticipated level of technical effort will vary based on each offeror's scientific and technical approach. Offerors must propose reasonable costs commensurate with their chosen approach. Please see Section 5.A.5 of the BAA for specific information.	2/04/11
012	Page 8 of the BAA makes it clear that even if the same set of researchers apply for two different tasks under the program, their proposals need to be self-contained, in particular have no cross-reference to each other. On the other hand, it is to be expected that some of the basic techniques and algorithms could be applied to more than one task. For example some of the techniques used in TA1 could have also applications to TA3.2. This raises the following question. If the same set of researchers	Proposals must be self-contained, but may contain technical efforts that are duplicative across multiple proposals due to a subtask providing a capability common to multiple proposals. If multiple proposals are selected from the same offeror, funding of common subtasks will be addressed during contract negotiations.	2/04/11

#	Question	Answer	Date Posted
	have proposals for two or more of the BAA tasks, can the proposals have overlap in the proposed algorithms and in the research questions to be investigated? Clearly, the two proposals can be written in a self-contained way but they may include overlapping text and planning when describing some of the underlying techniques (even if these will be applied/implemented in different ways depending on the particular task). Is that acceptable?		
013	Can you rephrase requirement TA1-A6? Since the Server is the only provider of records in its database, in what sense is there a need for Server verifying the integrity of this data? Is the verification you intend to be performed by the Server on the database in the clear, as it is input to the Server by whatever data source populates the Server's database?	TA1-A6 is intended to provide the Server with the ability to verify the integrity of data stored at a Third Party.	2/04/11
014	Can you elucidate on topic TA1-P11: What search do you envision within the content of the document tree? Should the query contain the XML tag and the search happens only in the content which is marked with the proper XML tag? And what kind of search on the content do you need? Should the search query specify e.g. an XML tag t and a keyword k and the search returns all XML documents which contain a field marked with tag t whose content (parsed as a sequence of alphanumeric strings separated by spaces) includes keyword k?	The Government is seeking innovative approaches to data retrieval in contexts in which data is not stored in tables. The XML example is a common example of non-tabular data storage, but proposals may address other scenarios. Each proposal may specify the search capabilities that are supported by the chosen scientific and technical approach.	2/04/11
015	Is it possible to have access to any previous proposals submitted within any previous Automatic Privacy Protection (APP) calls?	Please see Section 5.C of the BAA for information on proposal retention. Access will not be given to proposals from this or any other IARPA program.	2/04/11

#	Question	Answer	Date Posted
016	How important is scientific excellence and the standing of the PI in the scientific community as an evaluation criterion?	Please see Section 5.A.4 of the BAA for information on how the qualifications of the PI and other key personnel will be evaluated. As stated in the first paragraph of Section 5.A, the five evaluation criteria are listed in descending order of importance.	2/04/11
017	For TA3.1, TA3.2, and TA3.3, the BAA uses a comparison system (such as Amazon SNS) as a standard for performance requirements. Should we infer any additional requirements from the comparison system? For example, for TA3.1, should we infer that the 8K limit on the size of published items in Amazon SNS applies to the system to be built under SPAR?	During Test and Evaluation, any operational limitations of the comparison system on the underlying data will be reflected as nearly as possible by test data processed by the prototype under evaluation, consistent with the need to maintain assurances of security and privacy during the efficiency measurements. Note that the specific features of any named example comparison system should not be viewed as requirements or acceptable limitations on the proposed research prototypes.	2/04/11
018	In TA 3.1, what is the distribution of the size of submitted items?	There is no size distribution specified for TA 3.1.	2/04/11
019	In TA 3.1, may we assume that there is no requirement to add new clients as the system is running?	That is correct. TA3.1-R5 requires that the Server add and delete subscriptions dynamically, but this does not necessarily imply that the set of Clients is equally dynamic.	2/04/11
020	In TA 3.1, is it permitted for the server to learn that two different clients have a common interest?	That is correct. This type of access pattern information is not required to be protected by the requirements enumerated in Table 9. On the other hand, the Server may not learn the common interest (TA3.1-R1), nor what items or if any items matched that common interest (TA3.1-R4).	2/04/11
021	In TA3.1, is it permitted for the server to learn the number of interests that a client has?	Yes, the requirements enumerated in Table 9 do not exclude this information from being learned by the Server.	2/04/11

#	Question	Answer	Date Posted
022	In TA 3.1, are the clients also the entities that submit new items for possible publication?	This is not a requirement of TA 3.1. The source of items processed by the Server is not specified.	2/04/11
023	In TA 3.1, does a client that submits an item related to interest X have to be subscribed to interest X?	The source of items processed by the Server is not specified, so it should not be assumed that they originate with system Clients. There is no required relation between the receipt of new items by the Server and the subscriptions of Clients.	2/04/11
024	In TA3.1, at what rate do clients add and delete subscriptions?	There are no specified bounds on the rate of addition and deletion of Client subscriptions that a proposed approach can support. With regard to Test and Evaluation, the efficiency metric defined in TA3.1-R7 does not require a dynamic number of Client subscriptions.	2/04/11
025	In TA3.1, requirement TA3.1-R4, what does it mean for a client to "access" an item? In a publish/subscribe system, a client expresses its interests through a subscription and the system publishes items to a client. Generally, a client doesn't have an explicit capability to initiate "access" to an item.	In TA3.1-R4, Client access means the ability to retrieve an item that matches its subscription. Publish/subscribe systems may or may not support a true “push” capability. Often a true “push” capability is simulated by a Client program that periodically polls the Server for new information. In this case, Client access refers to these repeated requests for updates. In a true “push” scenario, in which items are sent to the Client without its initiation, access would refer to these transfer events. If an individual item is transferred to a Client via push or pull more than once, then the Client may be able to infer when it is deleted by the Server. Under the assumption mentioned in TA3.1-R4, this is not possible, but any Third Party should not learn when and which items are deleted under this assumption.	2/04/11
026	In TA 3.1, are there any requirements related to benign failures, like network partitions or processor crashes and restarts?	There are no specific robustness requirements for TA 3.1.	2/04/11

#	Question	Answer	Date Posted
027	For TA3.1-R7, the first sentence seems to require that the throughput of the Server be at least ten times better (i.e. - not worse) than that of the comparison system. The second sentence then defines throughput as the mean elapsed time between pairwise publication and receipt. This definition seems to be the inverse of the typical meaning of throughput: it is measured in seconds / message rather than the usual messages / second. Is the throughput requirement that the Server have a mean elapsed time between pairwise publication and receipt not longer than ten times that of the comparison system, both under the specified load?	The point about throughput being the reciprocal of the defined metric is duly noted. However, the items per second metric is easily obtained from the values actually measured. Please note that the requirement is mean elapsed time between pairwise receipt by the Server of an item and subsequent publication of an item to the Client. That is, “receipt” does not refer to final transfer of an item to the Client.	2/04/11
028	Will a proposal that addresses all of the requirements for TA3.3 except for all of TA3.3-R7 be considered if it meets all of the other requirements?	All compliant proposals will be considered, but proposals that do not address one or more of the requirements for a technical area may not be favorably reviewed. See Section 5.A, Evaluation Criteria and 5.B, Review and Selection Process for more information on how IARPA will evaluate proposals.	2/11/11
029	Can you please provide more information on the query types described as TA1-P10 in Table 4? Examples of usage scenarios or of specific operations that the program would like to address in this query type would be beneficial to assess the appropriateness of potential solutions.	TA1-P10 does not define any specific type or range of query capabilities for searching relational databases consisting of multiple linked tables.	2/11/11
030	Can you please provide more information on the query types described as TA1-P11 in Table 4? Examples of usage scenarios or of specific operations that the program would like to address in this query type would be beneficial to assess the appropriateness of potential	TA1-P11 does not define any specific type or range of query capabilities for searching data that is not stored in tables; however, the requirement does target structured data rather than free text documents. An XML document tree is mentioned in TA1-P11 for illustrative purposes only -- offerors	2/11/11

#	Question	Answer	Date Posted
	solutions. In TA1-P11, do you require the ability to search arbitrary text documents?	may propose other types of structured data and specific query capabilities related to those structures.	
031	Lines 23-24 on page 12 of the BAA say: "A Server is any entity with access to record encryption keys or plain text records (not as a result of a query), or access to a subset of cipher texts in their original plain text order." Does this definition of server apply to technical area TA3.1?	The referenced definition on page 12 of the BAA is meant to apply to the case in which a database contains records that are ordered with respect to their content. The BAA has been amended to allow parties other than the Server in TA3.1, TA3.2, and TA3.3 to have access to cipher texts in original plain text order.	2/11/11
032	Page 13 of the BAA says "Proposals that minimize the number of explicit Third Parties are preferred, with a strong preference for innovative approaches to practical two-party protocols." Which one of these proposal structures is of greater interest to IARPA? 1) design and implementation of a single protocol that meets all requirements while minimizing the number of explicit third parties, or 2) design and implementation of multiple protocols with tradeoffs between the number of requirements met and the number of explicit third parties? Furthermore, page 14 of the BAA says "Proposals that provide security and privacy assurances in adversarial models other than HBC are desirable and will be reviewed more favorably than proposals that provide assurances only in the HBC model." Which one of these proposal structures is of greater interest to IARPA? 1) design and implementation of a single protocol that meets all requirements in the least restrictive adversarial model, or 2) design and implementation of multiple protocols with tradeoffs between the number of requirements met and how restrictive is the adversarial model?	All compliant proposals will be considered, but proposals that do not address one or more of the requirements for a technical area may not be favorably reviewed. See Section 5.A, Evaluation Criteria and 5.B, Review and Selection Process for precise information as to how IARPA will evaluate proposals	2/11/11

#	Question	Answer	Date Posted
033	The requirements for TA3.1 do not mention any specifics about the “publishers” -- how many publishers are there? How many events do they produce? Is a publisher an “entity” in the security sense, i.e., does it have a distinct cryptographic identity?	In TA3.1, the Server is the single publisher who receives items for potential publication from an unspecified number of sources who are not considered participants in the protocol. Third Party participants may be proposed to facilitate the secure, privacy-preserving transactions between the single Server/publisher and multiple Clients, but such third party participants will not be considered publishers. The number of items made available by the Server for publication to subscribed Clients is not specified, but TA3.1-R7 states a performance requirement that must be achieved for a system in which, among other load parameters, one item arrives each second. The Server may consist of multiple component subsystems, but page 12 of the BAA (Protocol Participants, second paragraph) notes that Server components are considered to be the same entity with respect to security requirements. This consideration means that all security requirements of “the Server” must be satisfied by any component of the Server. It is not required that all Server components must share the same cryptographic identity.	2/11/11
034	In TA3.1, what kind of system assumptions are made regarding the publishers: do they interact with client(s) off-line (e.g., at subscription time only)? Are they assumed to have a business relationship?	In TA3.1, the Server is the single publisher, although Third Party participants may be proposed. No relationship between any parties should be assumed other than the mutual desire to share data while maintaining the specified security and privacy assurances.	2/11/11
035	In TA3.1, what kind of non-collusion/trust assumptions are made about the Server? Is the Server assumed to not collude with Client(s)? Publishers(s)?	Pages 13-14 of the BAA (Adversarial Models) describe the non-collusion and trust assumptions for all technical areas. In particular, the security and privacy assurances must be maintained under an Honest-but-Curious model in which no participants collude.	2/11/11

#	Question	Answer	Date Posted
036	In TA 3.1, can there be more than one interest associated with a document? If so, for test and evaluation of TA3.1-R7 what is the distribution of the number of interests?	The SPAR BAA does not restrict the number of interests that may be associated with each item processed by the Server. In many publish/subscribe applications, it is certainly reasonable to assume that an item would be of interest to multiple Clients for different reasons, or possibly to no Client at a given moment in time. The metric defined in TA3.1-R7 does not specify the expected number of relevant subscriptions/interests per item. Prior to test and evaluation, the configuration of the test and evaluation environment other than the load parameters and their minimum values specified in TA3.1-R7 will be established and applied to both the comparison system and the prototypes under evaluation.	2/22/11
037	In TA 3.1, would it be more correct for us to assume (1) that documents and their associated interests arrive in plain text on the server by some unspecified means or (2) that there may be some document provider that is a party in the protocol and on which we can deploy code (e.g., to encrypt documents)?	The source of items processed by the Server is not specified, so the assurances of the protocol must not be based on the source of items. In some applications, a Server may process items generated by external parties, and in others, the Server may itself generate the items. SPAR BAA section 1.A.2 (Protocol Participants) states that a Server is any entity in the protocol that has access to plain text items (not as a result of a query). Therefore, a document provider that has access to plain text items (not as a result of a query) would be treated as a Server component, and the protocol must then be agnostic about how items "arrive" at the document provider.	2/22/11
038	For protocols that support multiple Clients through a common proxy, is the proxy considered a "Third Party" participant?	No. Entities that serve as a common (proxy) node for multiple Clients are not considered "Third Parties." They are a type of Client node. See SPAR BAA section 1.A.2. (Protocol Participants).	2/22/11
039	For protocols that support multiple Clients, is there a preference between protocols in which Clients	No. The SPAR BAA does not specify a preference between protocols that support multiple Clients using a proxy and	2/22/11

#	Question	Answer	Date Posted
	communicate with the Server through a proxy and protocols in which each Client communicates separately with the Server?	protocols that support multiple Clients in other ways, all other characteristics being equal.	
040	What are more precise definitions of (1.3.a) “direct,” (1.3.b) “indirect,” (1.3.c) “burdened,” and (1.3.d) “unburdened” rates for the purposes of this proposal (section 2, last three paragraphs)?	The BAA uses the terms “direct cost” and “indirect cost” as they are defined in the Federal Acquisition Regulation (FAR) 2.101. Section 4.B.2, subsection 2 (page 51) uses the term “burdened rate” to refer to the total cost that does not identify individual cost elements, and uses the term “unburdened rate” to refer to a cost breakdown that shows direct and indirect cost elements separately.	2/22/11