

Automated Low-Level Analysis and Description of Diverse Intelligence Video (ALADDIN) BAA Questions

#	Question	Answer	Date Posted
001	Will existing metadata extraction tools or any output data from such tools be provided by the Government?	Neither metadata extraction tools nor the output of such tools for the test collections will be provided by the Government. The Content Description Representation (CDR) must be generated in its entirety as specified in Section 1.A.3 using a complete set of extraction technologies provided by the Offeror.	07/09/10
002	Is the intent to create the CDR exclusively using the existing state-of-the-art in metadata extraction technology or will novel extraction research be supported?	Where feasible, the existing state-of-the-art in metadata extraction technology should be leveraged in generating the CDR. However, novel research in metadata extraction will be acceptable if it will result in improvements in speed, accuracy, robustness, and expressiveness that are critical to achieving the Program's goals and that maintain an effective balance between performance, complexity, and cost. Novel metadata extraction research must not include the technical areas called out in Section 1.A.4, "Research Activities that are not of interest to ALADDIN."	7/09/10
003	Is there a restriction on organizations participating in more than one proposing team?	Individuals or organizations may participate on more than one bidding team. However, individuals or organizations associated with multiple teams must take care not to commit more resources than can be effectively applied to the ALADDIN program should more than one of their teams be selected for funding. Specific content, contributions, communications, networking and team formations are the sole responsibility of the performers.	7/09/10
004	Will the program be building technology to process static test data sets or continuously streaming incoming data?	As per Footnote 2 in the BAA, "In an operational setting, a CDR is populated with new video clip entries over time as new clips are added to the collection. For ALADDIN research purposes, all test collections will be fixed in size prior to evaluation and the video clips in each of these collections are to be processed en masse to create the collection's CDR."	7/09/10

#	Question	Answer	Date Posted
005	Can highly-specialized hardware be used in creating the CDR?	Yes. However, as specified in Section 1.A.3, "proposals must include a detailed description of the [CDR] system architecture, and offerors should explain how their approach provides a proper balance of speed, system complexity, and cost. Proposals that fail to address system complexity and cost, or that propose architectures that are highly complex and/or costly, will not be favorably reviewed."	7/09/10
006	Is the program seeking to build operational systems that support distributed environments and multiple users? If so, are there bandwidth and latency requirements for the event agents?	The Program is performing research to create key core technical capabilities that will enable the development of advanced video search tools. It is not seeking to develop a large-scale multi-user/distributed operational application. As such, there are no bandwidth or multi-user response latency requirements. The proposed technology must support a single user processing one event query at a time on a COTS standard personal computing platform.	7/09/10
007	Can the systems be designed to incorporate query relevance feedback and user correction of extracted results to improve the system? Can the prototypes be designed so that users can add manual annotations to the results?	The creation of the CDR and the execution of the event agents must be fully automatic. Interactive event query formation is permitted. But, relevance feedback mechanisms may not be used to achieve the Program's performance goals. These results must be achieved through a single query/search iteration. Additional interactive functionality that improves system performance or usability may be researched in the development of the demonstrable prototypes. However, such functionality may not be used in implementing the primary MED and MER evaluations that will be used to measure progress against the Program's performance goals.	7/09/10
008	Can the CDR be updated manually or automatically during evaluation?	As described in the Milestones in Section 1.C.2, the CDR for each evaluation may be refined and re-generated prior to the start of that evaluation. Once an evaluation begins, the CDR must remain unmodified until the evaluation is completed.	7/09/10
009	Will ad hoc event tasks be semantically related to pre-specified event tasks?	Offerors should not assume that Ad Hoc event tasks will be related to Pre-Specified event tasks.	7/09/10

#	Question	Answer	Date Posted
010	If the query event appears in one of the source video clips but is not the main subject of the clip, or the event occupies only a portion of the clip, will the ALADDIN system be expected to detect the event in such a clip?	Yes.	7/09/10
011	We want to confirm our understanding of the speed performance metric (the RT measure) for Event Agent Execution in Table 2. Suppose there are 1000 hours of source video in the video clip test collection. Would the speed performance goal for Event Agent Execution in Option Year 4 therefore be 1 hour?	Yes.	7/09/10
012	Does the processing time for Event Agent Execution exclude the time needed to generate the event agent?	Yes.	7/09/10
013	Are there examples of the kinds of events that the Program will address? Will the nature of the events change in each option year?	See the NIST TRECVID MED 2010 website (http://www.nist.gov/itl/iad/mig/med10.cfm) for examples of the kinds of events. The nature of the events can be expected to change during the Program. The focus in each option year is on improving performance as per the goals given in Table 2, not on processing events with increasing difficulty.	7/09/10
014	Is the program seeking research that can provide an integrative approach to the technical challenge rather than disjointed research efforts on how to solve the constituent technical problems?	Yes.	7/09/10
015	Can universities or companies apply, or teams of both?	No specific organizational structure is prescribed. As described in Section 1.D, "Teaming", offerors should propose a carefully-selected team and effective team management approach to address the Program's technical challenges and	7/09/10

#	Question	Answer	Date Posted
		performance goals. Organizational eligibility requirements are specified in Section 3, "Eligibility Information".	
016	Can we assume that the audio in the majority of video clips in the test datasets will contain English speech?	No.	7/09/10
017	Should we expect that the content of the speech in a clip may be a distinguishing feature for event detection, or is it safe to assume that relevant events are observable and distinguishable from appropriate video analysis and sound classification alone?	Per the Program's definition of an event in Section 1.A.2, "High Level Overview", an event occurrence must be "directly observable" in a video clip. This does not preclude use of information in speech to detect an event occurrence. However, the detection of spoken references to non-observable events is out of the scope of the Program.	7/09/10
018	Do the brief biographical sketches of key personnel and significant contributors count towards the 30 page limit?	Yes. The only exceptions to the page limit are specified in the opening paragraph of Section 4.B.1, "Volume 1, Technical and Management Proposal {Limit of 30 pages}."	7/20/10
019	Are face identification and speaker identification technologies included in the category of biometrics which are outside the scope of the ALADDIN program?	Yes.	7/20/10
020	What video domains will be focused on in the Program? Are there examples of the kinds of video clips that will be used?	Per Section 1.A.1, "Background", the Program is focused on "'unconstrained' video clips produced by anyone who has a digital camera." Per Section 1.B.3, "Data Resources", each of the NIST test collections will contain "tens of thousands of video clips from a variety of sources and genres." See the NIST TRECVID MED Website (http://www.nist.gov/itl/iad/mig/med10.cfm) for examples.	7/20/10

#	Question	Answer	Date Posted
021	By specifying MPEG-4, we are assuming that you are referring to the variant of MPEG-4 defined by the Motion Imagery Standards Board (MISB). Is that correct?	Per Section 1.B.3, "Data Resources", NIST is distributing the video clip data for its TRECVID evaluations in the MPEG-4 industry standard that is appropriate to the Program's unconstrained domain.	7/20/10
022	What additional metadata will be provided for evaluation in the Program?	In addition to the information presented in the event queries as described in Section 1.A.2, "High Level Overview", only the test dataset video files will be provided as input in the TRECVID MED and MER evaluations. While these may contain embedded metadata, no other information or external metadata will be provided.	7/20/10
023	Can we assume that all of the video clips will be in the raw original form they were recorded in?	No. Each video clip may have been arbitrarily modified.	7/20/10
024	Are the video clips used as event query exemplars also part of the evaluation test dataset?	No, the event query video clip exemplars will be distinct from the evaluation test dataset.	7/20/10
025	What data will be used to populate the Content Description Representation (CDR)?	For the ALADDIN Program evaluations, the CDR must be populated with data as specified in Section 1.C.2 "Milestones". A performer may use other data in their own research, subject to the requirements and restrictions given in Section 1.B.3, "Data Resources."	7/20/10
026	When will the pre-specified event task queries be given to the performers?	As stated in Section 1.B.2, Subsection "Event Detection and Recounting Evaluations", "the video clip data collection and event queries for pre-specified event tasks will be released near the beginning of each evaluation cycle."	7/20/10