PROPOSER INFORMATION PAMPHLET (PIP)

FOR THE

Video Analysis and Content Extraction (VACE) Program - Phase III

Broad Agency Announcement (BAA)

BAA 06-01-MT

Disruptive Technology Office (DTO)
Fort George G. Meade, MD

TABLE OF CONTENTS

VACE Program
Phase III
BROAD AGENCY ANNOUNCEMENT (BAA)

PROPOSER INFORMATION PAMPHLET (PIP)

GENERAL

The information provided in this pamphlet, in addition to that provided in the Federal Business Opportunities (FedBizOps) Announcement, BAA 06-01-MT, constitutes a Broad Agency Announcement (BAA) as contemplated in FAR 6.102 (d) (2) (i).

All administrative correspondence and questions concerning this BAA must be directed, in writing, to the following administrative addresses:

Contracting Officers Representative: Mr. Lawrence H. Carter
Department of the Interior
National Business Center
Acquisition and Property Management Division, Southwest Branch
PO Box 12924
ATTN: BAA 06-01-MT (L. Carter)
Fort Huachuca, Arizona 85670-2924
Voice: 520-533-1213
Fax: 520-533-1600
Email: **Lawrence_H_Carter@nbc.gov**

Contracting Specialist: Ann Peine
Department of the Interior
National Business Center
Acquisition and Property Management Division, Southwest Branch
PO Box 12924
ATTN: BAA 06-01-MT (A. Peine)
Fort Huachuca, Arizona 85670-2924
Voice: 520-533-1063
Fax: 520-538-0415
Email: **Anna_G_Peine@nbc.gov**

Contracting Officer: Gloria Golden
Department of the Interior
National Business Center
Acquisition and Property Management Division, Southwest Branch
PO Box 12924
ATTN: BAA 06-01-MT (G. Golden)

Fort Huachuca, Arizona 85670-2024
Voice: 520-538-0418
Email: **Gloria_M_Golden@nbc.gov**

Internet Web Site: www.nbc.gov/solicit.html

The Department of the Interior, National Business Center, Acquisition Services Division, Southwest Branch, Fort Huachuca, intends to use electronic mail for most technical and administrative correspondence regarding this BAA. Technical and contractual questions shall include the originator's full name and return e-mail address in the text. Questions and answers will be posted to the solicitation home page.

Written requests for information concerning this BAA may be sent by, as follows:

By facsimile:

520-533-1600, addressed to ATTN: BAA 06-01-MT (BAA INFORMATION, L. Carter)

By Email:

Email: **Lawrence_H_Carter@nbc.gov**

By surface mail (USPS):

Department of the Interior
National Business Center
Acquisition Services Division, Southwest Branch
PO Box 12924
ATTN: BAA BAA 06-01-MT (BAA INFORMATION, L. Carter)
Fort Huachuca, Arizona 85670-2924

By overnight delivery service:

Department of the Interior
National Business Center
Acquisition Services Division, Southwest Branch
Building 22208, Corner of Auger and Adair Streets
ATTN: BAA 06-01-MT (BAA INFORMATION, L. Carter)
Fort Huachuca, Arizona 85670-6000
Voice: 520-533-1213

If e-mail is not available, please direct questions to one of the above addresses. These requests must include the name, address, phone number, and email address of a point of contact at the requesting organization.

1 INTRODUCTION

The Disruptive Technology Office (DTO), a United States Government entity, formerly called the Advanced Research and Development Activity (ARDA), is soliciting proposals for Phase III of the VACE (Video Analysis and Content Extraction) program.

1.1 Background

ARDA was a research organization that was jointly established in December 1998 by the U.S. Department of Defense (DoD) (http://dod.gov/) and the U.S. Intelligence Community (IC) (http://www.intelligence.gov/). ARDA was organizationally part of the National Security Agency and its mission was to incubate revolutionary, disruptive research and development (R&D) activities within the broad field of Information Technology. While its mission remains the same, ARDA, renamed the DTO, has become an organization under the Director of National Intelligence (DNI). The DNI, itself, was established on 21 April 2005 to serve as the head of the U.S. Intelligence Community. The DTO falls directly under the aegis of the Associate Director of National Intelligence for Science and Technology (ADNI/S&T).

1.2 Overview

The DTO's individual research programs are organized around broad research topics called research focuses. One of these research focus areas is Information Exploitation (InfoEx), and under InfoEx is the VACE program. In all cases, DTO's research focus areas and programs have been established and executed for the shared benefit of the DoD and IC. In order to satisfy this mission, DTO, in close cooperation with its DoD and IC partners, originates and manages advanced research and development (R&D) programs that:

- will have fundamental impact on satisfying future DoD and IC operational needs and influencing strategies;
- demand substantial, long-term venture investment to spur risk-taking;
- progress measurably toward mid-term and final goals; and
- take many forms and employ many delivery vehicles.

Representatives from the VACE program and a number of DoD, IC and Civil agencies have developed this BAA for VACE Phase III under DTO's guidance and direction. The Department of Interior, National Business Center, Acquisition Services Division, Southwest Branch, Fort Huachuca, AZ (hereinafter referred to as NBC), has agreed to issue this solicitation. The evaluation of proposals, selection of awardees, execution of the resulting contracts, and overall management of the proposals awarded from this BAA will be accomplished under the guidance and direction of the VACE program.

2 VACE PROGRAM GOALS

Phase III is the next 3-year effort in the 9-year VACE program. It will be initiated during the summer of 2006 and run through the summer of 2009. (see Figure 1 below)

The Government, through the VACE program, is seeking proposals for innovative, creative, high-risk research to achieve significant advancements in technologies and methods for ingesting, indexing, managing, accessing and understanding a large video corpus from multiple heterogeneous video data sources. It is the intent of this research to continue to advance the state-of-the-art in technologies and methods for advanced, automated and automatic video content extraction, intelligent content services, and underlying enabling technologies.

TENTATIVE SCHEDULE

12/01/05	Notice of intent in FedBizOps; register interest; build teaming base
12/15/05	Draft BAA posted; register for Bidders Brief
1/06/06	Comments/questions cutoff
1/20/06	Bidders Brief; planning virtual meeting
2/1/06	Final BAA Announcement
3/03/06	Proposals due to Government
4/15/06 – 4/30/06	Evaluation recommendations completed (depending on response)
6/30/06	Contract awards completed
6/30/08	BAA extension period
6/30/09	BAA completion

Figure 1

The heterogeneous video data source domains may include, but are not necessarily limited to the following:

- Broadcast News;
- Meetings/Conferences;
- Surveillance;
- Ground Reconnaissance (i.e. scenes of various indoor and outdoor activities involving people, vehicles and facilities shot from handheld cameras); and
- Unmanned Aerial Vehicles (UAVs).

VACE technologies will provide:

- significant improvement in indexing and retrieval performance for video data;
- autonomous video understanding; and
- development of advanced applications/processing functions for video querying/retrieval, browsing, monitoring, mining, and content-based routing.

The VACE Phase III research program will be focused on extending and enhancing the capabilities developed in Phase II, making them more robust and scalable. VACE Phase III will be more focused on developing capabilities that are specific to video – such as developing techniques that exploit the temporal and motion properties of video image sequences. All research programs in Phase III will illustrate clear advantages over the techniques developed in Phase II and develop measured successes above the research results of Phase II. Figure 2 displays a compilation of relevant technologies and their proven or expected maturity on a notional VACE timeline. In the figure progression from left to right of a technology bar illustrates study/completion of the simplest to the most difficult capability in that technology. Also, the figure groups the technologies into three research objective categories that are described in detail in Section 4.

VACE Phase III will be focused on the longest-range goals of video exploitation, such as the understanding of the video content and recognition of specific events. This will be enabled by the research successes of previous VACE phases.

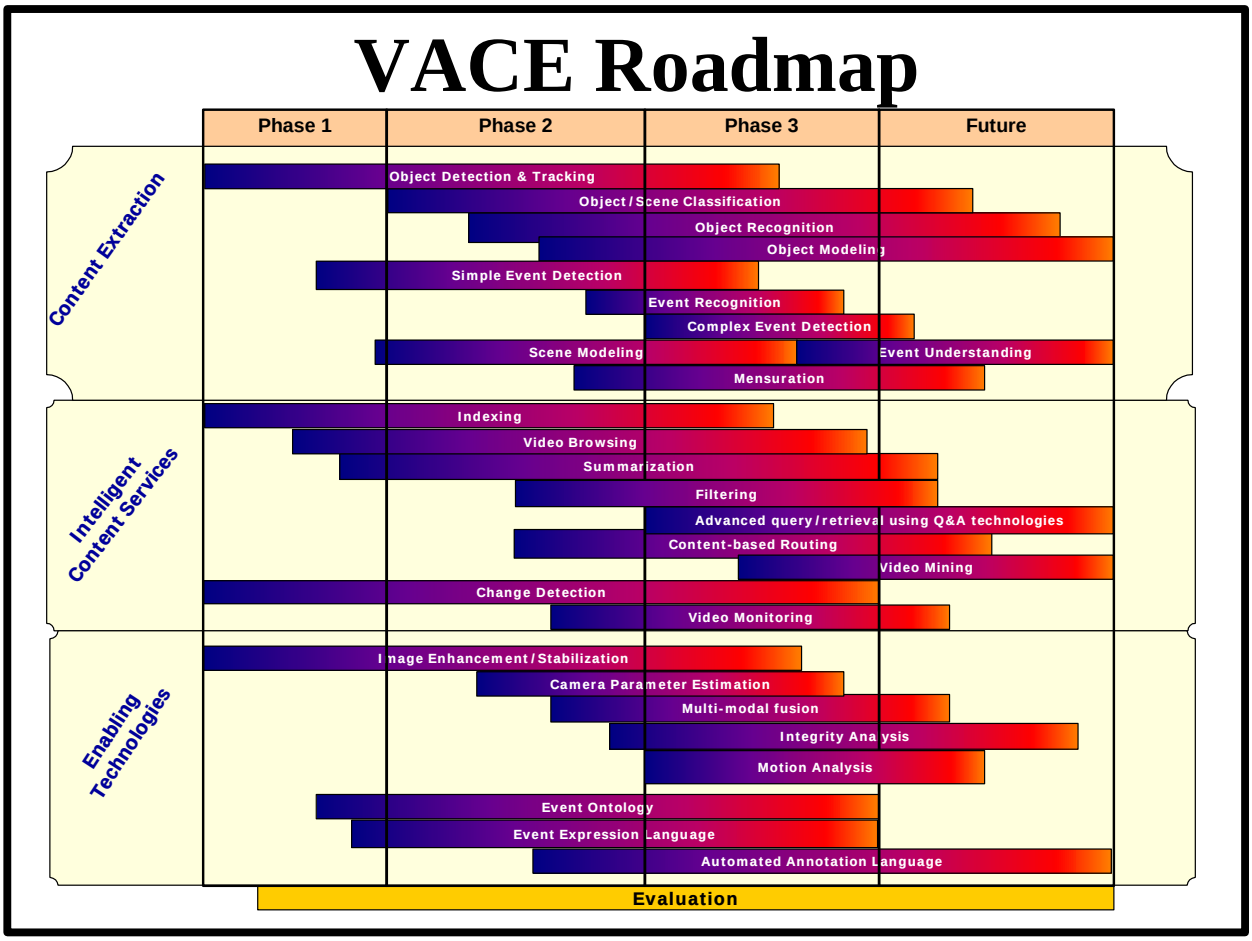

Figure 2

Two guiding precepts that drive this research and enable evaluating the level of overall performance are the ability to perform analyst tasks more accurately and more quickly than before. At the top level the primary metrics are accuracy and speed. Thus the

intent of this research is not only to enable capabilities that did not exist before, but also to allow the automation of video analysis activities to achieve a DNI mission to “analyze and disseminate accurate, timely, and objective intelligence…” Accordingly, placing the VACE mission in that context, it is recognized that there is a great disparity between the current state-of-the-art and necessary enabling technologies as notionally depicted in Figure 3 below:

Metric	Current	Objective
Accuracy	<<Human	>Human
Speed	>>Real time	<Real time

Figure 3

Thus the goal is not just to automate the video analysis processes and assist the human analyst, but also to perform these tasks with superhuman accuracy and speed.

3 SCOPE

3.1 Structure

In recognition of technological progress made in Phases I and II, the Government has decided to structure the Phase III effort into two tiers. The Tier 1 initiatives will address research into basic technology problems that have yet to be solved. Tier 2 initiatives will address end-to-end applications that will integrate basic technology applications currently available or those to be developed through research early-on in the initiative. Tier 1 shall address one or more of the technologies described in Figure 2 and Tier 2 initiatives shall address large component or system level solutions consistent with the system architecture depicted in Figure 4. Unlike the Phase II program, Tier 2 solutions must address applicability to more than one of the heterogeneous video data sources mentioned in Section 2 and described in more detail in Section 4.

3.2 System Scalability

All of the approaches, methods, architectures, algorithms and techniques that are incorporated into these emerging and evolving advanced video analysis systems must be scalable. To meet the challenge of ever growing, massive data volumes the technology product must have the potential to be periodically re-optimized for efficient ("real" time) execution or response time against increasing data volume requirements. A video analysis system that takes an inordinate amount of time to process a given

dataset will find little acceptance among analysts who are under constant and critical time pressures to provide “timely” intelligence.

Unfortunately, there is no upper limit in the foreseeable future for the magnitude of the data volume that the ultimate video analysis system will need to access. Consequently, Offerors are strongly encouraged to provide insight into how their resultant technology is to be developed with scalability and optimization in mind. In addition, Tier 2 Offerors shall consider system architecture methodologies and technologies that are flexible, configurable, and adaptable to multiple video domains and data formats that will reside within ever-increasing data repositories. As core algorithms are enhanced, these system architectures should facilitate integration of evolving video analysis and extraction technologies.

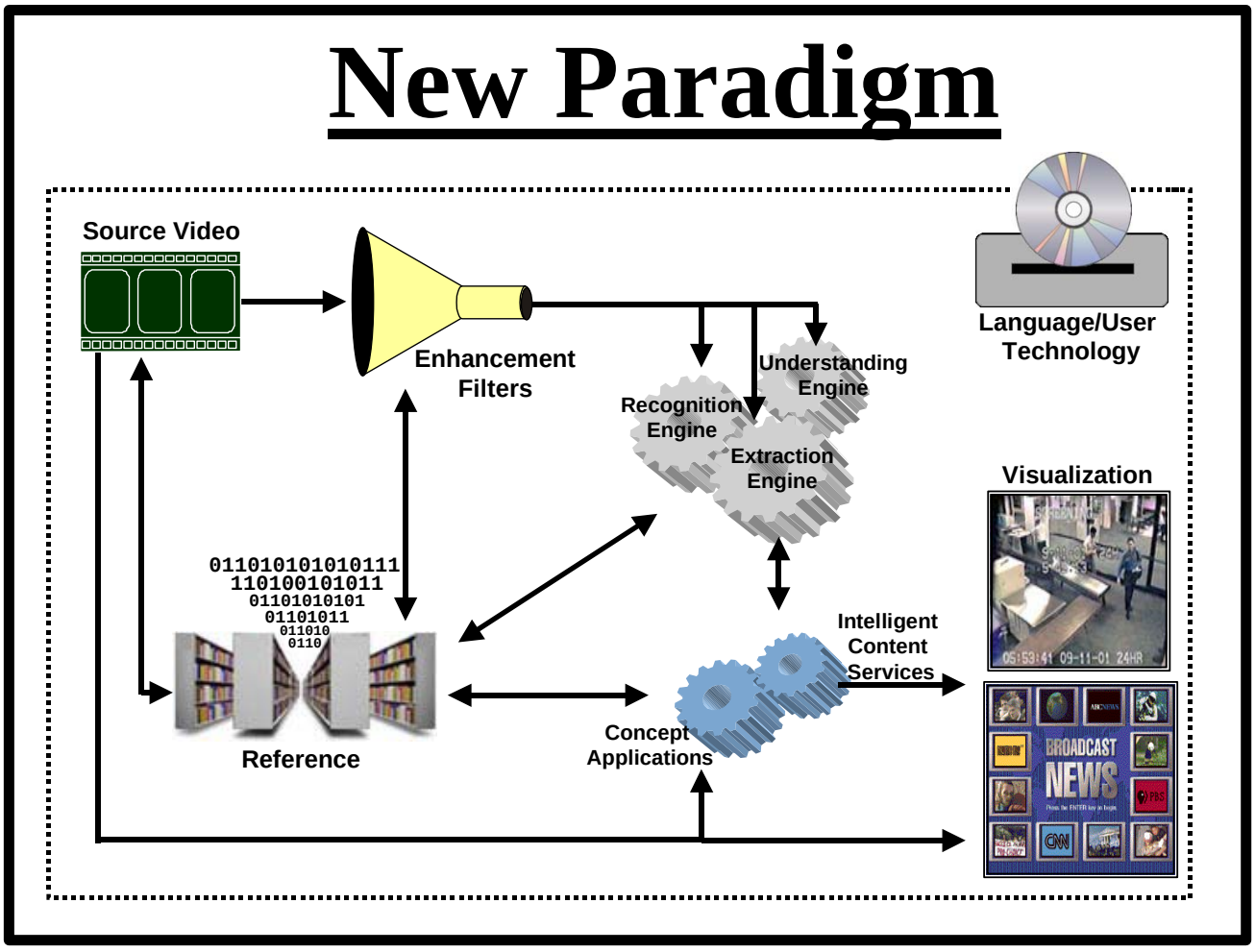

Figure 4

3.3 Eligibility
VACE program participation is open to all U.S and non-U.S. research and development organizations including:

- large and small businesses;
- academic and eligible non-profit and not-for-profit institutions; and
- collaborative ventures from mixed sources, as well as U.S. Federally Funded Research and Development Centers (FFRDCs).

All VACE program Phase II participants must submit proposals in response to this BAA to be considered for award under Phase III. All non-U.S. organizations are invited to submit proposals for Tier 1 initiatives but must team in a subcontract role with a U.S.-based organization for Tier 2 initiatives.

3.4 Period of Performance
Phase III shall have a period of performance of 24 months for Tier 1 and 36 months for Tier 2 contracts.

3.5 Total Program Funding
Funding for Phase III of the VACE program will utilize funds from FY 2006-2009 and is anticipated to total at least $ 30,000,000 over the BAA lifetime.

3.6 Individual Awards
The Government anticipates funding approximately 8 to 12 Tier 1, and 2 to 4 Tier 2 proposals at varying levels of effort.

- Tier 1 contract(s) will be incrementally funded for 2 years (base plus option year), with each year's level approximately $300,000 to $500,000 for a highly focused study.
- Tier 2 contract(s) will be incrementally funded for 3 years (base plus 2 option years) with each year’s level approximately $900,000 to $1,500,000 for a more comprehensive component/system level effort.

A single Offeror may submit independent proposals in either Tier 1 or Tier 2 or both, so long as the prime bidder in Tier 2 is a U.S company.

Upon selection for award, the contracting officer will determine fair and reasonable price, based on cost/price analysis.

3.7 Patent, Data Rights and DoD Specific Clauses
Offerors are hereby advised that any resultant contract will be subject to the following clauses:

- FAR 52-227-11 Patent Rights – Retention by Contractor (Short Form) (for small businesses) or FAR 52.227-12 Patent Rights – Retention by Contractor (Long Form (for big businesses);
- DFARS 252.227-7013 Patent Rights - Acquisition by the Government;
- DFARS 252.227-7014 Rights in Data – General; and
- DFARS 252.227-7015 Representation of Limited Rights and Restrictive Computer Software.

If the Offeror proposes to bring in data or software that has been developed at its own expense before this contract, then the Offerors are hereby advised that any resultant contract will be subject to the following clauses:

- DFARS 252.227-7037, Validation of Restrictive markings on Technical Data;
- DFARS 252.227-7013, Rights in Technical Data; and
- DFARS 252.227-7017, Identification and Assertion of Use, Release or Disclosure of Restrictions.

DoD Contracts will also include:

- DFARS 252.203-7001, Prohibition on Persons Convicted of Fraud or Other Defense Contract Related Felonies.

For Offerors who plan to subcontract with foreign companies/organizations, the following clauses will be applicable:

- DFARS 252.209-7001, Disclosure of Ownership or Control by the Government of a Terrorist Country;
- DFARS 252.209-7002, Disclosure of Ownership or Control by a Foreign Government;
- DFARS 252.209-7004, Subcontracting with Firms that are Owned or Controlled by the Government of a Terrorist Country.

The Offeror shall include in their proposal any data they propose to deliver under any resultant contract that will be subject to restricted rights.

3.8 Use of Existing COTS and/or GOTS Software Components

An Offeror may incorporate into their approach existing Commercial Off-the-Shelf (COTS) and/or Government Off-the-Shelf (GOTS) software components. Any knowledge of hardware platforms and/or operating system dependencies must be identified. This provides a framework that allows greater latitude in proposing innovative and revolutionary research in more focused areas.

CAUTION: Any software developed under VACE must NOT be so tightly coupled with any existing COTS and/or GOTS software that it becomes difficult or cost prohibitive for the Government to integrate it with other similar products. Offerors must clearly state

any plans for use of COTS/GOTS products, identify the COTS/GOTS products to be used, specify the cost and any assumptions about existing or emerging capabilities that they plan to use or on which their research depends.

4 RESEARCH GOALS

The VACE Phase III focus is on event extraction and understanding and those technologies that will enable those capabilities as shown in Figure 2, and realization of large component or system level solutions consistent with the system architecture depicted in Figure 4. This section defines what we determine as an event, and emphasizes the elements within events that are most desired by the VACE customers.

4.1 Video Research Environment

The VACE Phase III research environment consists of the relevant technologies shown in Figure 2 assembled into three logical research technology objective categories. These technologies and categories are described and defined in Section 4.2. The technologies operate on the five data source domains listed in Section 2. The data domains and some specific examples of DoD/IC user needs particular to those domains are described in Section 4.4.

4.2 Research Technology Objectives Categories

To better understand the following technology objective categories we provide the following definitions for clarity. For purposes of this BAA, an *event* is defined to be the occurrence of some form activity or action involving one or more entities. Webster's Dictionary defines the word *entity* as "something that exists as a particular or discrete unit." (i.e. low level event) Here entity differs from "object" or "person" in that an entity can be, but is not limited to a:

- collection of static objects or people (e.g. stockpile of weapons=armory, collection of gas pumps=gas station, etc.);
- scene comprised of objects or people and characterized by a simple action (e.g. multiple vehicles on moving paved surface = traffic scene, uniformed people marching in formation= military activity scene, etc.);
- one object or person characterized by simple action and scene setting (e.g. person collecting toll at toll booth=tollbooth attendant, person driving a taxi= taxi driver, etc.).

4.2.1 Content Extraction

Content Extraction embraces component-level technologies to extract a variety of metadata from the video signal into a richly annotated temporal log -- the goal being to create a completely symbolic representation of the activities, events, relationships, and important content in the video. These technologies range from low-level object detection and tracking technologies to high-level event-understanding technologies requiring information fusion and semantic interpretation.

The Content Extraction technologies as shown in Figure 2 are defined as follows:

- *Object Detection & Tracking*
    - Detection - The ability to determine the presence, spatial attributes (location/position), and pose of specific object types in a video scene. These techniques may also be used to further detect specific component parts of a larger composite object type, such as: the detection of eyes, nose and ears on a face; arms, legs, torso on a human body; and information about the specific observed pose of three-dimensional objects.
    - Tracking - In a scene where one or more entities exist, the ability to independently track the entities spatially as well as temporally (single or multi-camera environment) and determine the spatial attributes of specific unique objects throughout a video sequence. Tracking may include techniques that provide spatial coordinates information in the projected image domain, in a two-dimensional or three-dimensional model of the scene environment.
- *Object/Scene Classification* – The ability to categorize the object(s) in a scene beyond modeling. (e.g. If the object is a human, to be able to determine characteristics such as gender (male/female), age (infant/child/adult), and coloration (skin/hair); and if a scene, to be able to determine characteristics such as indoor/outdoor, time of day, meteorological conditions, scene content, etc.).
- *Object Recognition* – The ability to recognize an identified entity from a video clip in terms of the definitions given above.
- *Object Modeling* - The ability to build accurate models of specific objects that may be used for a variety of fusion-level tasks. Precise object models are often developed from training data to be used in motion analysis, object recognition, and mensuration.
- *Simple Event Detection* - The ability to detect the activities that occur within a video sequence, to include detecting the temporal bounds and classifying events.
- *Event Recognition* - The ability to recognize an event and associated sub-events and activities temporally in a video clip.
- *Complex Event Detection* - The ability to detect activities comprised of a set of several simple events that occur within a video sequence; to include a UAV video that detects at least 10 individuals (probably soldiers), most of whom are carrying rifles, rapidly load three SUV-type vehicles and depart down a road in a convoy formation.
- *Scene Modeling* – The ability to determine the scene geometry or atmospheric conditions of a video sequence. Scene models may be simple two-dimensional models, such as image mosaics, or more complicated three-dimensional models. Scene modeling algorithms may include techniques such as background extraction, depth calculations, and detection of planar surfaces.
- *Event Understanding* – The ability to put context to an event, once the event is recognized, such as assessing and concluding the event to be normal or abnormal so as to catalogue or call attention to it as appropriate.
- *Mensuration* - The ability to determine a number of different measurements associated with a given object. (e.g. What are the linear dimensions of a military

vehicle/piece of equipment? What is the volume of the cargo compartment of a truck?).

4.2.2 Intelligent Content Services

Intelligent Content Services embraces integrated technologies with user interfaces implemented in application-like systems to process the information created by metadata extraction and deliver information to a user. These include search, routing, alerting, summarization, storyboarding, time-lining, and other information delivery technologies.

The Intelligent Content Services technologies as shown in Figure 2 are defined as follows:

- *Indexing* - The process of preparing video for rapid retrieval based on the video's content. Automatically generate various summaries for each story segment including headlines, filmstrip storyboards, and shots. Central to indexing is metadata generation that succinctly summarizes the video content in terms of key words that allow data visualization over time and space (named entity or geographical location). Such metadata generation can be accomplished from many sources that include, but are not limited to: video soundtrack, video closed-captioning or overlaid text as well as the video content itself – faces, objects, scene content.
- *Video Browsing* - The ability to rapidly discover the presence of entities, assess entity importance, and correctly interpret the level of activity to draw analysts' attention to areas of the collection that are potentially interesting and relevant.
- *Summarization* - The ability to provide a compressed representation of the automatically extracted video content. The summary can be much more than just the mosaic presentation of a series of thumbnail images collected at video key frames. Event-level summarization will use recognized objects and events to compute an efficient representation of the video summary. Multi-modal presentation of the summary will combine audio, image, and text.
- *Filtering* - The process of sorting through all responses to a query in order to reduce and sort those responses to the most relevant subset. This process can be interactive or automatic. In the interactive mode the analyst may use multiple visualization techniques to rapidly prune out irrelevant responses. This interactive process may use learning techniques to further automate the filtering. The automatic mode may employ automatic techniques to rank order the responses and provide an accompanying explanation for the specific rank and again employ learning techniques to further enhance future filtering.
- *Advanced query/retrieval using Q&A technologies* - The process of decomposing user questions and identifying relevant information in video data and metadata to formulate a cohesive, domain related answer.
- *Content-based routing* - The ability to automate the transfer of new video sequences to analysts for whom the content is relevant, based upon: the occurrence and recognition of objects; co-occurrence of objects; occurrence of

events; text recognition in the video scene; or speech recognition, which could be used to route video data to analysts based on stated interests.

- *Video Mining* - The ability to enable analysts to identify similarities, patterns, trends, anomalies, or other associations in the content-based metadata that has been extracted from raw video data. In most cases the linkages that are established by video mining algorithms connect video sequences/images from different video segments and can operate in either a supervised or unsupervised manner.
- *Change Detection* - The ability to conclude something has changed in a scene that has been observed from two or more temporally separated collections of the scene from similar or dissimilar sensor modalities and describe/depict that change for the user.
- *Video Monitoring* - The ability to apply automation to focus attention on specific objects or events in cluttered scenes, recognizing events and objects that may have occurred in large volumes of prior surveillance data. Typically, human analysts often monitor surveillance video in real-time, viewing multiple data streams in raw form. These human analysts use motion and object recognition cues to focus their attention on specific objects and events.

4.2.3 Enabling Technologies

Enabling Technologies are technologies that enhance the performance of metadata extraction and intelligent content services technologies via pre-processing of the source video signal, and well as technologies that provide alternative renderings of the source video to the end-user. These include super-resolution, mosaicing, and geo-spatial mapping.

The Enabling Technologies as shown in Figure 2 are defined as follows:

- *Image Enhancement/Stabilization* - The ability to improve the entire image of a specific region/object within a video sequence by providing better spatial resolution and/or signal-to-noise ratio. This may be accomplished by: exploiting redundant information from multiple video frames; using domain knowledge or constraints imposed by the object type; and/or exploiting knowledge of compression or image formation processes.
- *Camera Parameter Estimation* - The ability to determine the camera position, focal length, and other related camera parameters based solely on information contained within the video stream without any a prior knowledge.
- *Multi-modal fusion* - The ability to combine multiple types of data processing to extract salient information about whether to do object detection, object recognition, or event understanding of any given scene. (e.g. I see that there is a map of an amusement park captured in a video. Given information from an old map and photographs of street names, what can you understand about the amusement park?)

- *Integrity Analysis* - The goal of video integrity analysis is to verify the authenticity of encoded video sequences. An individual video is examined for signs of manipulation or tampering.
- *Motion Analysis* - The ability to characterize/classify the movement of detected and tracked objects that are modeled as one or more rigid objects. The motion of a single object may be analyzed to determine one of a class of possible activities or actions. Motions may consist of a single motion or a sequence of motions.
- *Event Ontology* - The ability to provide ontology or other type of hierarchy that captures the interrelationships that exist between individual events. Viewed from the top down, this ontology or hierarchy should describe the decomposition of more complex events into their more elementary sub-components. Viewed from the bottom up, this ontology or hierarchy should describe the manner in which simple elements may be combined to form more complex, comprehensive, and descriptive events. Ontology or hierarchy will need to include logical, spatial, and temporal constraints between individual events.
- *Event Expression Language* - The ability to provide a formal language for representing and encoding automatically extracted video events. The language needs to be both expressive and expansive to simplify and enhance the human analysts' interaction with automatically extracted video content data. Specifically, the language should be able to capture common video events, describe interactions between objects at an appropriate level of detail, and express logical, spatial, and temporal relationships.
- *Automated Annotation Language* - A language to represent the metadata output of multiple detectors that can be used for further processing (i.e. Summarization). The method used to interface metadata content extraction technologies with intelligent content services.

4.3 VACE General Research Objectives

VACE research objectives are specific to the needs expressed by the DoD and IC. Consistent with previous VACE solicitations, automated content extraction from video is still a key objective. In this phase, emphasis is placed on not only the raw ability of a Proposer's algorithm, but the processing speed (baseline 30 fps) and the amount of data the algorithm can handle. Each proposal shall list anticipated processing speeds and improvement expectations for the life of the project. Likewise, each Proposer shall list the expected amount of data their algorithm can handle with respect to processing speeds. The performance should be conformant to National Television System Committee (NTSC) and Phase Alternating Line (PAL) specifications[1].

4.4 VACE Specific Research Objectives

The more specific research requirements are listed in data source domain specific areas in the sub-sections below. Though some of these research objectives require basic vision technologies like object detection and tracking as a foundation to solving

[1] NTSC - U.S. video standard; PAL - European video standard (see http://www.webopedia.com/TERM/P/PAL.html for definition)

the problem, the Proposer must emphasize the efficacy of their approach as it addresses solving the research areas listed in section 4.2.

> *NOTE (1):* This BAA intentionally specifies only what types of information are needed but does not specify the process(s) required to obtain them.
>
> *NOTE (2):* There intentionally are cases where the same question(s) to be answered is listed under more than one data domain. This is done for emphasis but does not preclude the Offeror from proposing to answer the question over more than one data domain.

4.4.1 Broadcast News Exploitation

Televised broadcast news, and more specifically foreign news television broadcasts, have a wealth of information. As an arm of the foreign Government they frequently present positions and propaganda for homeland consumption that may not be available in other media, particularly if print media is managed separately. Video broadcasts frequently present images of objects, people, activities or events with a voice tracks that excel over reports in other forms. Independent news broadcasts can provide significant insight into how a country is operating (i.e the health of democratic processes including free speech and dissent, the relative strengths of various political forces, economic conditions, and many other important issues). The problems posed by analyzing these broadcasts start with the vast volume of world broadcasts, followed by limited strategies to get analysts the information they need in a timely fashion. Sometimes this is a pull (when an analyst knows what they want or can specify it in an archive query or by browsing). Sometimes it is a push (when something occurs, can be identified, and is routed to the proper analyst). Because there are many layers of information in video news broadcasts and a number of quality or reliability factors to be considered, there must be filtering processes that help the analyst navigate the video quickly.

The interest in Broadcast News technology falls into two categories: (1) that which enables ingesting, indexing, managing, accessing news video; and (2) that which enhances content understanding. As seen in Figure 2, the technologies associated with (1) have been studied for some time and have reached a certain level of maturity. Emphasis in these technologies is in the ability to speed up the associated processes. The primary interest of Broadcast News is to understand content and thus the emphasis of this BAA is on tool and algorithm development that focus on:

- detecting shots;
- the format of the scenes (indoor, outdoor, studio, split-screen, still image, pure text, illustrations, maps, interview, round-table discussion, debate, speech to large or small audience, etc);
- the speaker, when speaking;
- the speaker by name, when captioned or by automatic speaker recognition;
- the presence and number of major objects in a shot;
- the presence and number of people in the shot, even when partially or temporally obscured;

- the identities and relative locations of people in a scene when identified by caption, narration, in-scene text, re-appearance in a fixed position during the broadcast, or by biometrics;
- identification of companies, countries, and other entities based on recognized logos, national flags, or other graphical representations in or on the scene;
- identification of a forum or event (such as a world trade summit or state visit of a leader);
- geographic or geo-political identification of the location of a scene when identified by caption, narrative, or other identifying feature;
- identification of gestures or gait in association with detection of a human;
- extraction of an object model involving multiple dimensions of processing;
- the motion of objects or people in a scene with relative or ordinal direction and speed estimation;
- object recognition within an object class, such as military objects, associated with mensuration of essential or major dimensions;
- detection of specific event classes within the ontology/taxonomy of events proposed by the Offeror;
- recognition of a specified subset of video events within the detected video event classes;
- mapping between the audio story of the foreign news broadcast and the detected/recognized video events; and/or
- articulation of inconsistencies between detected or recognized events with the audio story.

4.4.2 Meetings/Conferences Data Domain

Meeting/Conference video covers indoor meetings in a variety of venue, format, and topic. The venues may range from small informal settings to large conference halls seating hundreds of people. The format of the meetings will range from informal gatherings to structured educational presentations and press conferences. These meetings may involve group discussions, visual presentations, and question/answer sessions. The topic of the meetings will range from technical presentations to group planning exercises. There may be multiple cameras for each scene, but the cameras typically will be stationary. Cameras will have the capability to pan, tilt and zoom. Audio may or may not be available for all video. In cases where audio is present, the conversation and text may be in a foreign language.

In the meeting/conference data the research should focus on answering the following questions:

Who was involved in the meeting?

- Can we disambiguate instances and represent all human entities in a meeting?
- Can we recognize specific individuals from observed biometric measurements?
- Can we recognize specific types of clothing?
- Can we read text on nametags and nameplates?

- Can we disambiguate specific participants from appearance or context (clothing, gestures, mannerisms, etc)?

What is the relationship between participants?

- Which participant is the leader and which are subordinates?
- Which participants have direct interactions during the meeting?
- What types of interactions do the participants engage in (sidebar conversations, angered exchanges, looking at)?

What is discussed in the meeting?

- Can text on whiteboard, projection screen, or paper be detected and converted to searchable form?
- Can visual cues be used to determine temporal boundaries on topics that are discussed?

What events took place?

- Can we locate and recognize specific simple events in meetings (such as call-to-order, voting, discussion, adjournment, etc)?
- Can we locate and recognize more complicated events?
- At what point were decisions made?
- What were the decisions that were made?

What is the context of the meeting?

- What is the style of the room (i.e. auditorium, lecture room, meeting room, office, etc)?
- Can we automatically create a 3-D model of the room and its objects?
- Can we recognize specific meeting rooms?
- What is the meeting style (committee, lecture, demonstration, etc)?

4.4.3 Surveillance

Surveillance embodies collecting continuous video data to monitor events that take place in the field of view. The source of the video collection may be cameras that are located at various strategic locations, situated at various heights above the ground and positioned indoor or outdoors (or both). These cameras may or may not have an overlapping field of view, may be fixed or mobile (pan and zoom), and may collect data from various portions of the electromagnetic spectrum.

In surveillance, VACE Phase III emphasizes the detection, recognition and understanding of video events associated with the video output of surveillance video cameras. The goals are:

- to accurately and reliably monitor the ingress and egress of people, vehicles, and people with small or large objects;
- to track object movement within the field of vision of the surveillance camera system;
- to identify (within the duration of a scene) that a person or object is within view of the system, the “best” views of the object or person to aid in a variety of recognition schemes;
- to understand these patterns of activity and movement well enough to distinguish between normal and abnormal activity;
- to do the above robustly in the presence of complex event patterns, deception, and imagery that is degraded by environmental effects, compression, or sensor decay; and
- to research and develop tailorable interfaces for real time surveillance; rapid search, retrieval, and retention of critical sub-events; and for retrospective search to identify trends and long-term activity patterns.

VACE Phase III addresses the problem of adequately distinguishing normally occurring actions from events that are of interest. Current security and monitoring systems generally issue alerts based primarily on simple motion detection, which can rapidly generate an unacceptable level of false positive alarms.

Systems with intelligence guided by computer vision have great potential, but they still are weak in many respects. Activities that should be candidates for alerts, to the extent that they can even be segmented and recognized from atomic actions, may be obscured by such issues as:

- a high density of occluding activities that are of no interest;
- occlusions that may or may not be intentional;
- complex and unfamiliar patterns that can be mistaken for normal activity; and/or
- temporally displaced actions that make a composite event difficult to isolate.

In this request for proposals, the Government is not seeking proposals for surveillance hardware advances, but recognizes that hardware advances may be enablers for computer vision advances for which proposals are desired. Individual cameras may exhibit imaging capabilities that are not common to all cameras in a system to accommodate location-specific requirements. Human operators may control cameras (pan/tilt/zoom) to observe activities or to anticipate them. Some cameras may image in infrared or other spectral ranges. 360-degree cameras, such as an iMove camera (http://www.imoveinc.com/), may be used to obtain the maximum view from a single camera location. High-Definition cameras may be used to maintain the larger context and still get good detail on small components such as personal characteristics for biometrics or numbers on license plates.

Numerous research and development opportunities exist for research into human recognition, which is an important component of an intelligent surveillance system. VACE Phase III Tier 2 will not support new work in this area, however systems should anticipate integration of some form of biometrics where appropriate. Some level of demonstration may be appropriate using existing recognition systems.

Our vision of the future for these systems is to:

- recognize and discriminate simple and complex events on high semantic levels to make the technology easily accessible to end-users/operators;
- detect conditions for undesirable events by distinguishing patterns of activity that are normal versus abnormal, or that indicate elevated risk;
- detect undesirable events as they happen;
- provide tools to rapidly analyze a situation after an event has occurred;
- provide tools to rapidly review vast quantities of data for long term analysis of activities at a location;
- rapidly integrate multimedia (i.e. other sources of video information that may circumstantially be available after an event to improve situational awareness and understanding including sources such as the “ground reconnaissance” topic in this request or independent surveillance system data that may not be fully calibrated).

In VACE Phase III the focus shall be on increasing the quantity and variety of events of interest that can be recognized and doing so efficiently on video feeds of long duration and with numerous cameras.

4.4.4 Ground Reconnaissance

Ground Reconnaissance video contains scenes of various indoor and outdoor activities that involve objects (e.g. people, vehicles, buildings, etc.), places (e.g. facilities), and activities all shot from handheld cameras. The collections are expected to range from tourist camcorder video to that collected by troops/convoys in a combat zone. The use of these videos is to collect information of intelligence value in the form of: identifying objects; verifying locations; looking for changes over time; developing 3-D texture models of locations, facilities, bases, or cities for purposes of modeling; and simulation for situational awareness, training, and mission planning and rehearsal. These videos may or may not be taken with the intention of collecting intelligence information and may even be video that is provided through a third party.

Since the potential intelligence benefits span a wide range of activities, the requirements listed below are not all inclusive but are meant to lay out basic needs and stimulate ideas:

- Situation Awareness and Context -- determine the setting of a video. Examples of questions to be answered are the following:

- Is it indoors or out?
- What are the lighting levels?
- Is there a way to determine if a wall in a scene is interior or exterior?
  - Can you describe the interior scene?
- What season is it?
- What are the meteorological conditions?
- What is the sun's position in the sky?
- Can the cardinal axes of the scene be derived?
- Is it day or night?
- What is the phase of the moon?
- What kind of scene is represented (location and/or activity)?
  - place name or geographic coordinates
  - city/rural
  - land/water
  - street/building/facility
  - a battle/conflict
  - missile launch
  - gathering/parade
- Can objects and faces be detected? Recognized?
- How old is the video?
  - Exact date and time?
- If audio, what is the language?
  - Can a voice be recognized?

- Scale and Mensuration -- an ability to accurately determine scene scale, across the scene (both foreground and background) such that sizes of objects and people can be made, relative angles and distances between items of interest can be determined, or constant scale mosaics or 3-D models of the scene can be constructed. These geometric measurements should be accomplished with or without varying levels of a prior knowledge. Examples of questions to be answered are the following:
  - What are the sizes of objects (both stationary and moving)?
    - Can the biometrics data of a human(s) be determined?
    - Can the size (and possibly temperature) of plumes from missiles be determined?
    - Can you read serial numbers or other information on an object?
  - If indoors can the size of windows and/or doors in a room be measured?
    - What are the dimensions of the room?
  - What is the distance between objects?
  - What is the velocity of moving objects?
    - vehicles, etc.
  - Can a map or mosaic of the scene be generated?
    - uncontrolled scale/controlled scale
  - Can a 3-D virtual model of the scene be derived?

- Integrity Analysis -- determine how the video was captured, modified, stored, etc. Examples of questions to be answered are the following:

    - Has the video been edited?
    - Has the video been converted to/from European PAL format?
    - Has a video been taped over?
    - Can you determine who has shot video footage based on the shakiness of the footage?
    - Can you determine where edits are in a video clip? (not shot boundaries, but splices that are intended not to be discovered)
    - Can it be determined if video-quality changes from scene to scene in the same clip?
    - Can the resolution be enhanced?

4.4.5 UAV (Unmanned Aerial Vehicles)

UAV motion imagery data represents video recorded from airborne platforms. UAVs come in all sorts of shapes and sizes. Some fly at high altitude and some fly at low altitude. The video may be shot from a wide variety of downward-look angles. They have varying sensor suites that consist of cameras that are tuned to collect data from various portions of the electromagnetic spectrum (sometimes simultaneously) and can rotate, pan, and zoom at an operator's discretion. Sometimes the cameras may have video stabilization and/or auto-focus.

The UAV mission profile may be one of reconnaissance or surveillance. The distinction is that reconnaissance may or may not have a specific set of targets and may be searching for a specific activity (a general geographic area or a specific facility) or just for targets of opportunity. Conversely, surveillance typically has a known target(s) such as: an object (e.g. building, vehicle); an event or activity; or the motion of people, vehicles, or other moving objects. The UAV data is linked to a ground station for immediate or subsequent analysis. The metadata associated with the video imagery again varies from system to system. In many cases it is very sparse and limited to collection parameters and approximate location. It is the function of the analysis process to use the video in conjunction with the metadata to extract intelligence information as well as embellish the metadata. These two activities are of interest for the Phase III. A description of the analysis process and its needs follow in the ensuing paragraphs.

Exploitation and analysis of airborne *motion imagery* currently takes place in support of DoD/IC missions. (It should be noted that in the IC, UAV video source data is referred to as motion imagery and so that name is carried out in this subsection.) For first-phase analysis, motion imagery is received as live streams and is clipped into segments that are integrated into web-based intelligence reports. Motion imagery is also archived into a library, which is available to analysts to search and retrieve data. Analysts are then able to used archived data for use in second-phase and third-phase analysis.

First-phase analysts screen live streams of video, clip segments of interest and generate web-based reports that include clips related to mission objectives as well as analysts comments. First-phase reporting addresses immediate needs of a mission or intelligence issue, and the turn-around for products is typically near real-time (a few minutes).

Second-phase analysis involves performing additional analysis and merging related intelligence data using motion imagery as a reference point. Part of this process is product generation, and could include such products as mosaics, super-resolved video or still images, color correction, detail enhancement, stabilization, geo-referencing and geo-registration. Reporting and products are typically generated over days or weeks (generally longer than a week).

Third-phase, data mining is performed by the analyst in order to assess patterns, collect related events and enable integrity analysis. Advanced data mining supported by well-defined data models would foster detailed and finely filtered database queries. In-depth analysis is supported by the use of complex algorithms for more robust content extraction. Analysis at this level tends to take longer than first-phase or second-phase analysis as analysts utilize a range of data and technology to help solve an intelligence problem. Reporting and products are typically generated over days or weeks (generally longer than a week).

The principal focus of this BAA is on the tools necessary for late second-phase and third-phase analysis. In this UAV data domain the tools or algorithms that are developed should enable the analyst answer at least the following questions:

General:

- What activities or events are occurring in the video?
- Is there text in the video and what does it say?
- Are there people present in the video?
- Are people getting together or are they dispersing?
- Where are they coming from and where do they end?
- What's their mode of dressing?
- Are people moving quickly or slowly?
- Are they carrying anything and if so is it a weapon or package (e.g. gun, rifle, shoulder missile, IED etc.)?
- Can distinction be made between adults and children, men and women?
- Can a leader be identified in a group base?

Vehicles:

- What types of vehicles are in the video - heavy equipment, cars, trucks, tanks, tractor-trailers, motorcycles, boats, ships, rafts, etc?
- Are vehicles loaded or are they being on/off loaded?
- What are the vehicles dynamics and interactions (meetings/stopping)?
- What surface are vehicles being driven on; paved road, dirty road, river, sea, water, and ice?

- What happen at intersections?
- What are the interactions between the occupants in a vehicle?
- Did vehicle occupants leave vehicle and board another?
- How long vehicle stopped before resuming the journey?
- Did the vehicle disappear into a tunnel, garage, bushes etc?
- Can it be determined that efforts are being made to conceal vehicle activities?

It should be noted that analogous questions could be associated to speedboats on waterways.

- Where is the speedboat - river, sea, harbor, etc?

The interoperability of data and products through the intelligence cycle and between different tools and technologies is enabled by standards, as defined in the Motion Imagery Standards Profile (MISP), for motion imagery, metadata and products. Proposers are encouraged to consider these standards in their research plans.

In addition to the challenge of processing video from a moving camera, the research should: consider the potential of leveraging existing and evolving research to better enable exploitation; the use of video with other sources of data; strive for real-time or near-real time processing; and design the resulting tool(s) for integration with other technologies and/or migration toward a commercial tool.

4.4.6 Other Technologies

This category is provided for the submission of proposals that address problems in video research that may not be described above but that the Proposer believes important to achieving video requirements described in this BAA. In addition, there may be a time during the life of this BAA that requirements may change or the technology focus may change necessitating additional or further enhanced technologies beyond those described above. Though proposals that solely address this “other” category will not have priority in the funding schema, there is always the potential of a selection based on the innovation of the Proposer.

4.5 Evaluation Overview

Proposals must demonstrate a clear understanding of the purpose and role of evaluation in the research and effective integration of evaluation into the research and development process. As such, the technical proposals shall include an outline of the required evaluation tasks and how objectively measured significant progress will be shown during the course of the research. The evaluation tasks should cover all key components of the research. If these tasks are comprehended by current/planned DTO sanctioned evaluations, the proposal should indicate this. If the Proposer intends to participate in a non-DTO-sponsored evaluation to satisfy this requirement, the proposal should contain a description of the evaluation program and how the results of the evaluation will be determined and published. If new evaluation tasks and infrastructure are required to support the research, these requirements should be clearly specified using the VACE Exploratory Evaluation task template provided in Appendix B.

4.5.1 Evaluation Plans

The evaluation plans should have a clear description of the tasks to be evaluated, data to be employed, annotations to be created, schedule, and metrics to be used. These tasks should be clearly mapped to the VACE technical objectives in Section 4.

Where possible, generally accepted DTO or Vision Community accuracy/error metrics should be specified for evaluation of these tasks. For fully automated tasks, these metrics should be objective and repeatable using fixed test material. For tasks requiring a user in the loop, standard usability metrics should be specified including user satisfaction, time-to-completion and/or steps-to-completion.

Since the goal of the VACE program is the creation of accurate technologies with faster-than-real-time performance, all proposals must address how algorithmic efficiency will be addressed. Therefore, algorithm speed is to be measured in addition to accuracy for all evaluation tasks. For completely automated tasks, processing speed shall be reported in number of times real-time as specified in Appendix C.

Evaluation tasks requiring user interactions should report standard usability metrics such as time and steps to completion. For all evaluation tasks, the planned hardware/operating system environment is to be specified. Where possible the evaluation task algorithms should be implemented on a single COTS CPU so that they may be reasonably compared. The particular video frame rate, resolution, color depth, compression method, etc. used should be specified when reporting this information.

All evaluation tasks must be baselined for accuracy and speed so that progress may be measured. Where possible, proposals should include a current baseline for both accuracy and processing speed for existing technology – preferably using data sets from previous DTO-sanctioned evaluations or other public evaluations and generally-accepted metrics. If the research objective is so novel such that no such baselining is possible at the time the proposal is written, the proposals should specify an estimated baseline and a plan for the measurement of an actual baseline on an initial system within the first 6 months of the contract.

The proposal must also specifically address how progress can/will be measured during the course of the research and what the goals will be for the end of the research. The performance of awarded contracts will be measured against these goals.

4.5.2 Evaluation Budgeting

A minimum of 10% of the cost budget for every proposal shall be reserved for participation in evaluations. This budgeting shall include the labor and travel/logistics expenses to participate in these evaluations. The evaluation budgeting shall include the following activities:

- participation in evaluation planning activities including teleconferences, meetings, and email exchanges towards the development of the guidelines documents, formats, and reference annotations that will govern the evaluations;

- the adaptation of research algorithms to provide the required output for processing by the evaluation software;
- implementation of pilot “dry-run” evaluations intended to shake down the evaluation process;
- implementation of developmental intrinsic (internally run) evaluations
- implementation of formal extrinsic (public) evaluations; and
- participation in evaluation workshops and conferences (in addition to the bi-annual VACE PI workshops).

This budget shall NOT include the development of algorithmic capabilities, even if they are intrinsic to the requirements of the evaluation. Nor should it include hardware/software purchased to develop those capabilities. These shall be considered to be core research activities, and the funding for these activities shall be specified within the core research component of the contract proposals.

4.5.3 Data Acquisition Objectives

Data will be identified, obtained, appropriately annotated, and distributed to support demonstrations and evaluations for the VACE program. The types and amounts of data are outlined in Appendix A. In the video domain, such publicly available source data and annotations are extremely difficult and expensive to obtain given: the Government’s stringent requirements for experimental design and Institutional Review Board (IRB) approvals; the complexity of the required data collection systems; the large amount of storage required to compile, edit, and distribute the collected video; and the lack of standards and software tools for annotation. Great efforts are being made to address the lack of available research corpora in the VACE program. The reality is that these resources are slow in coming.

In Phase III, the VACE program will include a strong and tightly integrated evaluation infrastructure to address these deficiencies. This infrastructure will be part of an international collaboration in terms of evaluation task management and data resources to leverage the available resources. The goal of this effort will be to:

- develop an international community which can bring greater weight to bear on the technical challenges;
- share source data, annotations, evaluation methodologies, and annotation and evaluation tools;
- focus the community on important key research challenges of interest across programs – both domestic and abroad; and
- share evaluation results, analyses, and knowledge regarding algorithmic improvements via technical exchanges at evaluation workshops.

Each proposal shall identify data their research requires beyond that listed, and answer the questions found in the data section of Appendix B.

Data annotation, and tools/techniques used for data annotation will be provided by DTO through governmental and contractual sources. The proper ground truth annotation of the data is only possible through interaction with the researcher. Each Proposer shall have at least one representative participate in all discussions associated with annotation of data required for their specified research evaluations.

4.5.4 Evaluation Workshops

Periodic public evaluations have brought the greater research community (domestic and international participants) to bear on the problem of interest, and post-evaluation technical workshops have provided a venue for knowledge exchange. This approach has been extremely effective both in improving the performance of the technology and informing the Government about the technical capabilities and deficiencies. More importantly, it has helped to build a community with the critical mass needed to solve difficult problems. This community extends beyond the boundaries of any one research program or geo-political region. As such, some of these workshops may be in international venues.

In order to maximally leverage the evaluation infrastructure that can be supported by the VACE program and to increase the number of researchers focusing on evaluation tasks of interest, the VACE program will be increasing its participation in cross-program/multi-national open evaluations. The VACE program will collaborate with other domestic and international Vision research programs to support evaluations of common interest. This approach will make a greater variety of important evaluation tasks and research corpora available for use by the VACE researchers. The results of these evaluations will be presented at designated international evaluation workshops. To make the evaluation workshops associated with these multi-national evaluations equally accessible across programs, up to half of them will be held in foreign venues – most likely in Europe. The VACE participants will be expected to present their work in these evaluations at these workshops and publish papers on their results for the proceedings of those workshops.

These workshops will provide a forum for researchers to interact on specific program-level issues that arise in the process of Phase III, such as coordination of evaluations or integration issues. These 1- or 2-day forums ensure agreement on protocols, metrics, data sets and timelines. Some workshops may utilize video teleconference to two or three remote locations to reduce travel costs from remote participants.

It is expected that the Principal Investigator of each awarded contract and each significant subcontract (or their designated senior technical representative) will attend each of the workshops for evaluations they participated in. The contractors are strongly encouraged to include members of their research staff (graduate students, post-doctorals, and even junior researchers) in these workshops so that they can provide in-depth details of their particular work and learn about the work of others in the area. The proposals must include plans and budgets for participation in these workshops. These cost estimates should assume at least one such workshop per year for each major technical area addressed in the proposal (e.g., search, extraction, recognition, etc.).

5 GOVERNMENT-DEVELOPED TECHNOLOGY READINESS ASSESSMENT

5.1 General

Although the focus of the VACE program is to identify and research unsolved technical problems, it is important to remember that the ultimate goal of the program is to develop innovative technologies that will perform autonomous data extraction, annotation, and analysis of large volumes of video content to enable an intelligence analyst to dramatically reduce his/her production time, and moving them from a rapid-prototyping environment to an environment in which the emerging technologies can be reviewed and evaluated is an important step on the way to full system development.

DTO established a Technology Readiness Assessment (TRA) process during VACE Phase II to facilitate the movement of emergent systems out of the laboratory and onto a preliminary staging platform for review, demonstration, evaluation, and integration. This methodology will continue to be followed in Phase III. The TRA Agent (TRAA) will be used to guide research prototype products though various testing, evaluation and demonstration levels.

A method employed by the TRA to evaluate progress along that maturity pathway is the DoD Technology Readiness Level (TRL)[2] assessment and certification process. The TRAA will report TRLs for each VACE technology initiative at the component, sub-system or system level to the VACE program manager (PM). At the research outset each technology initiative will be evaluated as to its current TRL and a projected completion maturity level shall be established by which progress will be assessed.

The ultimate TRA goal is to head off and correct problems that might occur if items are deployed in DoD and IC operational environments too soon in an immature or untested state. This methodology attempts to balance the needs for a standardized process against the needs for case-by-case assessment of, and support for, individual contractor research initiatives. Throughout the duration of the VACE Phase III research, prototype systems will be transformed and integrated into this environment.

As Video Analysis systems have grown progressively more complex, it has become increasingly more important for the research community to standardize component interactions so that components and sub-systems could be more widely used. Collaboration on designing an environment in which these components can be variously and interestingly integrated and tested is encouraged. The TRA offers a pathway through which technologies could then be evaluated in both classified and unclassified analytic environments within the VACE program's sponsoring DoD and IC agencies.

The intent of the VACE program is to use the results of these on-going integration efforts for a wide variety of purposes to include, but are not limited to, the following:

[2] See Appendix D for complete TRL definitions

- capability demonstrations using open source as well as operationally available video;
- performing end-to-end, large scale evaluations of the effectiveness and added efficiency of a combined, integrated video analysis system in a controlled environment;
- developing early, operational prototypes designed to solicit feedback from real, operational users, attempting to find meaningful intelligence to real operational problems;
- identifying potential, operational spin-offs that are ready for larger scale, technology insertion;
- providing additional insights into the successes and failures of the VACE program so that subsequent research can be meaningfully steered in directions that have the potential for higher payoffs; and
- evaluating and track the maturity level of the component, subsystem or system.

This TRA environment will provide the opportunity to determine which components, subsystems, and full-systems could be successfully integrated and/or combined in ways that will demonstrate even greater potential for application against solving existing operational problems across a broad spectrum within the DoD and IC.

5.2 Participation in Technology Transition

Offerors are expected to include in their proposal a plan for moving their technology into the TRA environment. This plan will identify various milestones during the life of the project when components, subsystems, and full-systems (with upgrades) can be moved into the TRA. VACE's TRAA, in concert with government researchers and technologists from across the DTO and/or the sponsoring DoD and IC agencies, will work directly with the VACE contractors to facilitate the migration of these prototype systems into the preliminary staging location. Early movement of prototype elements to the TRA is strongly advised, especially for those Offerors who previously participated in the VACE program. The TRAA will help produce appropriate application program interfaces (APIs) and other integration standards/conditions, and resolve problems that may arise during the integration process.

The ultimate success of the VACE program will be assessed upon the degree to which major research advances can be quickly, widely, and effectively transitioned into practical solutions that satisfy multiple, critically important DoD and IC operational problems. The TRA environment provides an intermediate step for determining the effectiveness, robustness, interoperability and extensibility of a research assumption before a system is fully developed for operational use. It also provides a venue in which to develop a "user pull" that is key in migrating a technology product from the laboratory to an operational environment.

6 ADMINISTRATIVE RESPONSIBILITIES IN PHASE III

Program participants are expected to assume responsibility for administration of their projects and comply with contractual as well as program requirements for reporting, attendance at program workshops and symposia, and availability for site visits.

6.1 Review Process

The VACE program will conduct frequent reviews of both the overall program progress and the individual contract performance through informal 1-day project kickoffs, 1-day project reviews, semi-annual 3-day program-level reviews, and periodic 1-2 day evaluation/integration workshops.

6.1.1 Program Reviews

There will be a Program Review (kickoff) at the beginning of Phase III after the successful Offerors are under contract. The focus of this review will be:

- the objective(s), schedule, and roadmap of Phase III as stated by the VACE program;
- an introduction of all the successful participants and their teams;
- an overview of each project by the participant's Principal Investigator as was described in the successful proposal;
- a discussion of planned evaluations; and
- interaction among the various participating teams.

There will be semi-annual Program Reviews that will focus on:

- technical aspects of the program;
- program-level evaluation and data issues; and
- facilitating open technical exchanges, interaction, and sharing between the various program participants.

These reviews (or workshops) will be held in a conference environment at a location selected by the Government. They will include technical presentations and demonstrations by each contractor, during which the contractor will openly describe the technical aspects of their research, results of evaluations conducted, and progress/successes/failures that have occurred as part of their funded research.

6.1.2 Project Reviews

These reviews are held at a site proposed by the contractor and approved by the Government. The project's Contracting Officer Technical Representative (COTR) and other members of the VACE program, as well as any interested, potential DoD and IC users will attend to discuss project specific technical and administrative issues. Appropriate progress reports, details of successes, and issues and contributions to the program goals will be expected at these reviews. It is anticipated that at least one

project review for each contract will occur during each period between program-level reviews (workshops).

6.2 Technical Symposia

At the Government's discretion the Government may host a periodic symposium in an on-going effort to merge the developing research insights and system results. The program will continue to ask researchers to share, in the environment of a small symposium, their underlying methodologies, system or component architectures and designs, and process flows. The goal is to provide an opportunity to all community participants to obtain insight into how a system/component is designed and functions, how it addresses a particular video challenge, and how successful the research efforts have been.

6.3 Project Status Accountability

Teams will provide the following to their Contracting Officer Representatives (CORs) and the VACE program management staff:

- The contractor shall prepare, and present at the project kickoff meeting, a Project Plan that establishes the goals and objectives for each task, and activities required to reach them. The contractor shall include a Work Breakdown Structure (WBS) to the second level. The contractor shall identify all reviews and deliverables and shall delineate the schedule and acceptance plan associated with each completed task.
- The contractor shall provide formal Bi-Monthly Status Reports that include an overview of technical progress against the plan, its management, and cost data.
- The contractor shall conduct Quarterly Reviews and Progress Demonstrations for the sponsor. These shall be presented at the contractor's facility or at VACE program semi-annual workshops.
- The contractor shall prepare a draft 60 days prior to the end of the period of performance; and a final technical report NLT 60 day's post-period of performance describing the methods used and the results of the research.
- The contractor shall prepare appropriate software documentation for the code, a user guide, and an administration manual. The contractor shall demonstrate and informally instruct the sponsor or their designate in the operation of the software.
- The contractor shall provide a final (as-published) copy of all publications that result from work conducted in this project.

In addition, periodic reviews by COTRs and the VACE PM will continue in Phase III as described under "Project Reviews" above.

7 PROPOSAL PREPARATION INSTRUCTIONS

This announcement is an expression of interest only and does not commit the Government to pay for proposal preparation costs. The cost of preparing proposals in response to this BAA is not considered an allowable direct charge to any resulting contract or to any other contract. However, it may be an allowable expense to normal bid and proposal indirect costs as specified in FAR 31.205-18. If a subcontract(s) with a Federally Funded Research and Development Center (FFRDC) is proposed, Offerors are reminded of the limitations in their use (see FAR 35.017) and must provide documentation in the proposal that work is not otherwise available from the private sector.

Discussions with any of the points of contact shall not constitute a commitment by the Government to subsequently fund or award any proposed effort. Only Contracting Officers are legally authorized to commit the Government.

It is anticipated that Phase III awards will be made by NBC. Some DoD and IC Agencies may require additional clauses in their contracts, which may need to be negotiated at the time of contract award. These clauses will be Agency-specific but may include such items as “This Agency may prohibit non-U.S. citizens from all or certain aspects of the work to be performed under any resulting contract. The fact that the Offeror intends the use of non-U.S. citizens on any resulting contract will not necessarily disqualify the company from consideration nor may the non-U.S. citizens be prohibited from working on some or all aspects of any resultant contract. “

Offerors may submit proposals covering up to and including a 24-month or 36-month effort. Offerors can discuss a vision of what future research efforts might include after the 24-month or 36-month period, but no formal proposals will be accepted for any period beyond 36 months.

7.1 General

Proposals shall consist of three volumes: Volume 1 - Technical / Management Details; Volume 2 – Additional Reference Information; and Volume 3 - Cost. The page format shall be 12 point or larger type, single-spaced, one-inch margins, single sided, 8.5 by 11-inch pages. Unnecessarily elaborate brochures or presentations beyond that sufficient to present a complete and effective proposal are not desired. Proposals exceeding the maximum total length **WILL NOT** be considered.

Electronic Proposal Format: Electronic proposals shall be made using Microsoft Word and Excel for Windows applications (compatible with Windows 2000 or Windows XP). CDs shall be clearly labeled, referencing BAA 06-01-MT, marked with the Offeror’s organization and proposal title (short title recommended). Hard copy and electronic media must be submitted together. If using Microsoft Word, embed any Microsoft PowerPoint graphics used. Microsoft Word documents, with graphics as separate files, are **NOT** acceptable. If video clips are submitted they must be compatible with the following media players and CODEC delivered with standard Windows XP:

- QuickTime 6.5
- Windows 7
- Real 10

Volumes 1, 2 and 3 must each be contained within a single electronic file, i.e., a single file containing all of Volume 1, a second single file containing all of Volume 2 and a third single file containing all of Volume 3. All electronic media must be verified virus-free by using up-to-date, reputable virus detection utility, such as Norton or McAfee anti-virus software, and so noted on the diskette or disk label.

Number of Copies: Offerors shall submit:

- one (1) original version paper copy of all three volumes of each proposal with original signatures
- two (2) electronic copy containing all three volumes in Microsoft Word for Windows (Microsoft Excel for any spreadsheet submissions) format on CD-ROM by the closing date

Information or data contained in a full proposal deemed proprietary by the Offeror shall be clearly marked**.** The Offeror must mark the proposal with a protective legend in accordance with FAR Part 15.6, Use and Disclosure of Data, (modified to permit release to outside evaluators retained by either ARDA or the Department of the Interior, National Business Center, Acquisition Support Division, Southwest Branch, Fort Huachuca) if protection is desired for proprietary or confidential information.

7.2 Volume 1 – Technical And Management Details

7.2.1 Cover Sheet

The Cover Sheet provided at Appendix E – Part I, and Appendix E – Part II of this document shall be completed by the Offeror and submitted with the proposal. All information requested must be provided. The CAGE, DUNS/CEC, and TIN codes provided shall be those of the Offeror and not of the principal place of performance, if the two are different.

7.2.2 Length

Volume 1 shall be no longer than 40 pages for Tier 2 and 25 pages for Tier 1 in length. This page limitation includes all information (i.e., figures, tables, graphics, charts, indices, photographs, foldouts, appendices, key personnel, etc.) but does not include the Cover Sheets, provided that they contain no substantive text. Foldouts will be counted as two pages and must be no larger than 11 x 17 inches. Offerors are encouraged to submit concise, but descriptive, technical proposals.

7.2.3 Content

The content of Volume 1 shall consist of two parts: namely, Part 1: Summary of Proposal and Part 2: Detailed Proposal Information.

7.2.3.1 Part I: Summary of Proposal

This section shall provide an overview of the proposed work, as well as introduce associated technical and management issues. This section shall provide a technical description of the project in sufficient detail to provide clear, quantifiable technical objectives and a technical approach with a project schedule showing definite decision points and endpoints. In a manner of the offertory's choosing, this Part shall provide a succinct description of the uniqueness and benefits of the proposed project. Offerors must address:

A. A brief overview of the innovative claims for the proposed research. (Include in this part all proprietary claims to results, prototypes, intellectual property, or systems supporting and/or necessary for the use of the research, results, and/or prototype. If there are no proprietary claims, this shall be stated.

B. Brief summary of the technical rationale, technical approach, and constructive plans for accomplishment of technical goals.

C. Schedule and milestones for the proposed research, including overall estimates of cost for each task. A one-page graphic illustration that depicts major milestones of the proposed effort arrayed against the proposed time and cost estimates must be included.

D. A summary of the deliverables associated with the proposed research.

E. A clearly defined organizational chart of all anticipated program participants with brief biographical sketches of key personnel and significant contributors, their roles (including role of Principal Investigator) and their level of effort in each year (calendar year or academic / summer year) of the program. A chart, such as the following, is suggested.

Participants	Org	Role	Year 1	Year 2
John Doe	ABC University	Key Personnel / PI	25%	35%
Peter Fillmore	ABC University	Key Personnel		
Mary Smith	ABC University	Significant Contributor	50%	50%
Doctoral Candidate 1	ABC University	Contributor	25%	25%
Doctoral Candidate 2	ABC University	Contributor	40%	40%
Graduate Assistant 1	ABC University	Contributor	50%	50%
Abigail Stone	XYZ Co.	Key Personnel	25%	25%
Ronald Johnson	XYZ Co.	Significant Contributor	40%	50%
Graduate Assistant 1	XYZ Co.	Contributor	25%	50%

7.2.3.2 Part II: Detailed Proposal Information

This part shall provide the detailed, in-depth discussion of the proposed research. Specific attention must be given to addressing both the risks and payoffs of the proposed research making it desirable to pursue. This Part shall provide:

- innovative claims for the proposed research. This is the centerpiece of the proposal and shall succinctly describe the unique proposed contribution.

- detailed technical rationale supporting the approach employed. The technical rationale shall clearly show why the proposed technical approach is expected to achieve the stated purpose within the proposed cost and time schedule. The rationale shall also describe the rationale for the claims and deliverable products outlined elsewhere in the proposal and show how past / current performance justifies an award in this technical area.

- Statement of Work (SOW), describing the effort's scope, the specific tasks to be performed and their associated schedules. At a minimum, the statement of work shall consist of the following sections:

    1. **Scope—a statement as to what the SOW covers: objectives and goals and major milestones for the effort. Key elements are task development and deliverables.**
    2. **Task / technical requirements—a description of tasks, representing the work to be performed, developed in an orderly progression and in enough detail to establish the feasibility of accomplishing the overall program goals. The overall effort shall be grouped into major tasks and identified in a work breakdown structure (WBS)-like numbering system. NOTE: Proposed costs shall have a one-to-one correlation to this reporting structure, which shall be depicted in the cost volume.**

- a description of the results, products, transferable technology and an expected technology transfer path must be included.

- comparison with other on-going research, highlighting the uniqueness of the proposed effort/approach and differences between the proposed effort and current state-of the-art clearly stated. Identify the advantages and disadvantages of the proposed work with respect to potential alternate approaches.

- discussion of Offeror's previous accomplishments / work in this or closely related research areas.

- description of the facilities that would be used for the proposed effort.

- if any portion of the research is based on the use of Government-owned resources of any type, the Offeror shall specifically identify the property or other resource required, the date the property or resource is required, the duration of the requirement, the source from which the resource will be obtained, if known, and the impact on the research if the resource cannot be

provided. If no Government-furnished property is required for conduct of the proposed research, this section shall consist of a statement to that effect.

- detailed description of the support, including formal teaming agreements, required to execute the Offeror's proposal. Discussion of teaming relationships shall include the programmatic relationship of team members; the unique capabilities and relevant accomplishments and concise summary of qualifications of all team members (key personnel and significant contributors), with information about their major sources of support and commitments of their time; the task responsibilities of team members; the teaming strategy among the team members; and the management approach for the team. Full resumes / curriculum vitae of key personnel and significant contributors shall be included in Volume 2 (Additional Reference Information) of the proposal.

- a summary of any proprietary claims to results, prototypes, or systems The Offeror shall submit a separate list of all technical data or computer software that will be furnished to the Government with other than unlimited rights in accordance with DFARS 252.227-7017, Identification and Assertion of Use, Release or Disclosure Restrictions. ***All*** VACE contractors will be required to provide deliverables (software and documentation) for integration with other VACE program contractor's products for use in TRA evaluations and demonstrations in an end-to-end simulated operational environment. (See Section 5.4 for more information about the TRA.)

- description of how progress toward completion of their research goals will be measured, including a description of the evaluations to be performed, a schedule of implementation and type of report to be prepared.

- identification and description of anticipated data sources to be utilized in pursuit of the project research goals.

- summary of a plan, schedule and process for participation in the VACE TRA.

7.3 Volume 2: Additional Reference Information

7.3.1 Cover Sheet

The Cover Sheet provided at Attachment 1 of this document shall be completed by the Offeror and submitted with the proposal. Include the cover sheet at the beginning of the file containing Volume 2. All information requested must be provided. The CAGE, DUNS / CEC, and TIN codes provided shall be those of the Offeror and not of the principal place of performance, if the two are different.

7.3.2 Length

No absolute page limit is set for Volume 2. However each individual resume / curriculum vitae provided in this Volume may not exceed 5 pages.

7.3.3 Content
This volume shall include:

- a brief bibliography (annotated, if desired) of relevant technical papers and research notes (published and unpublished) which document the technical ideas on which the proposal is based. This material will be used at the discretion of evaluators, to enhance their understanding of relevant related work. (No page limit is set.);

- individual resumes / curriculum vitae of key personnel and significant contributors.

7.4 Volume 3: Cost Information.

7.4.1 Cover Sheet
The Proposal Pricing Sheet at Appendix E Part II shall be completed and submitted with each offer.

7.4.2 Length
No page limit for Volume 3 has been established.

7.4.3 Content
The content of Volume 3 shall consist of two parts: namely, Part 1: Summary of Costs and Part 2: Supporting Cost and Pricing Information.

7.4.3.1 Part 1: **Summary of Costs**
This section shall include:

- a one-page cost and fee summary correlating with the milestones summary and schedule portion of the technical proposal;

- detailed cost summary shall be provided for the entire program, supported by breakdowns, as follow:
    - **3 By tasks / subtasks, correlated to Volume 1, Statement of Work Task Technical Requirements.**
    - **4 Labor hours by labor category**
    - **5 Materials by vendor quotes and purchase history**
    - **6 Subcontractors and consultants**
    - **7 Travel**
    - **8 Other direct and indirect costs**

7.4.3.2 Part 2: Supporting Cost and Pricing Information
This part shall include supporting cost and pricing information in sufficient detail to substantiate the summary cost estimates in Part 2 above. Costs for subcontracts having 20% or more of the total value of the work must be substantiated to the same

level of detail as the costs of the Offeror. The contracting office will also require a detail of ODCs for travel, equipment and supplies, etc. Offerors are requested to provide a copy of any indirect rate agreement, cost accounting system and purchasing system approvals, and a copy of latest CASB approval letter. Additionally, all Offerors must complete Representations and Certifications online at www.orca.gov.

Note: CDRL shall include a requirement for bi-monthly cost reporting

7.5 Handling of Proposals; Use of Consultants
All proposals shall be handled as competitive information; contents will be disclosed only for the purposes of evaluation and only to members of the source selection panel.

The Government intends to use consultants and/or contractors to assist in handling, managing, and evaluating the proposals. These personnel will have signed, and will be subject to, the terms and conditions of non-disclosure agreements. By submission of its proposal, an Offeror agrees that its proposal information may be disclosed to the aforementioned personnel for the limited purposes stated above. However, only the Government will make final award determinations under this BAA.

7.6 Proposal Submission
Proposals are due on or before [30 days after release date of BAA – adjust beyond 30 days to avoid weekends and Mondays.] Proposals will be submitted to the following address:

Department of the Interior
National Business Center
Acquisition Support Division, Southwest Branch
Post Office Box 12924
ATTN: BAA 06-01-MT (G. Golden or Ann Peine)
Fort Huachuca, Arizona, 85670-2924

Proposals must be submitted in accordance with the requirements and procedures identified in the BAA and this PIP. To be considered, full proposals (in original, and electronic media) must be received. All Offerors shall complete Representations and Certifications online at www.orca.gov. **The initial cutoff date for receipt of proposals is March 3, 2006. Subsequent cutoff dates shall be established in 60-day increments.** Proposals received after the cutoff dates and times will be evaluated in the next evaluation cycle.

For overnight package delivery, proposals shall be addressed:

Department of the Interior
National Business Center
Acquisition Support Division, Southwest Branch
ATTN: BAA 06-01-MT (G. Golden or Ann Peine)

Bldg. 22208, Corner of Auger and Adair Avenues
Fort Huachuca, Arizona 85613-6000

- **Proposals submitted by fax or electronic mail are not acceptable and WILL NOT BE CONSIDERED.**
- **Proposals and/or proposal modifications received after the BAA closing date will be handled IAW FAR 15.208.**
- **Proposals not adhering to the form and format required by this BAA WILL NOT BE CONSIDERED.**

7.7 Proposal Selection Criteria
Proposals will be selected through a technical / scientific / business decision process with technical and scientific considerations being most important. Proposals will not be evaluated against each other since they are not submitted in accordance with a common work statement. Evaluations will be performed using the following criteria listed in descending order of relative importance. Each of these criteria is more fully described in Sections 7.8.1 through 7.8.3. Proposals unresponsive to the Technical and research areas addressed in the BAA will not be fully evaluated and will not be considered for award.

1. Overall scientific and technical merit;
2. Potential contribution and relevance to VACE R&D program and Research Goals (See Sections 2 and 4); and
3. Cost and Schedule realism

7.8 Evaluation Guidance For VACE Program Phase III
This section contains a description of each of the factors to be evaluated from technical and cost perspectives, and the associated adjectival ratings. Each VACE program Phase III reviewer will rate each of the three evaluation criteria identified in Section 7.7 above as falling into one of four adjectival ratings.

Excellent: The Offeror presents a proposal that is clear and concise on all aspects covered by this criteria and offers excellent approaches and methods which, when executed, will fully meet and very frequently exceed all elements contained in this criteria.

Very Good: The Offeror presents a proposal that is clear and concise on almost all aspects covered by this criteria and offers very good approaches and methods which, when executed, will meet and may exceed some elements contained in this criteria.

Satisfactory: The Offeror presents a proposal that is understandable on all aspects covered by this criteria and offers satisfactory approaches and methods which, when executed, will minimally satisfy all elements and may occasionally exceed a few elements contained in this criteria.

Unsatisfactory: The Offeror presents a proposal that is not completely understandable in every aspect covered by this criteria and offers one or more unsatisfactory approaches and methods which, when executed, will not minimally satisfy one or more elements contained in this criteria.

After all reviewer assessments have been collected the VACE program Phase III Evaluation Panel will meet for the purpose of assigning a final adjectival rating against each of the three evaluation criteria to each proposal. The VACE program Phase III Evaluation Panel will use these final review adjectival ratings in its award recommendations.

7.8.1 Overall Scientific and Technical Merit

This evaluation criterion covers three subfactors that are listed in descending order of relative importance. Each reviewer will consider their assessment of each subfactor individually and then jointly, when assigning one of the four adjectival ratings listed above to these criteria.

Scientific Principles. The assessment of this subfactor will include the extent to which the Offeror's proposal bases its proposed approach on sound scientific principles, building on the foundations of previous technical contributions, or presenting a well-justified premise. Of particular interest is the extent to which the Offeror presents a technical approach to the development of VACE program capabilities that addresses a clearly stated relevant technical problem appropriate for advanced research and development, with a clear path for proving or disproving a highly relevant premise. Also important is the extent to which the proposed research represents extremely innovative, unique and creative approach to developing capabilities for the VACE program.

Strategy. The assessment of this subfactor will include the extent to which the Offeror's Statement of Work clearly describes and outlines a strategy, which if successfully implemented will result in the Offeror reaching the proposal's stated goals and objectives, clearly outlines task and technical work requirements in an orderly progression and in enough detail to establish the feasibility of accomplishing the overall program goals and clearly identifies frequent, measurable milestones. This subfactor also includes the extent and manner in which the Offeror's proposal describes its strategy for participation in evaluation and testbed activities.

Experience. The assessment of this subfactor will include the extent to which the Offeror's proposal describes the experience and capabilities of proposed research staff, and the appropriateness of that staff given the proposed research approach. Also of importance is the track record of the Offeror and the Offeror's organization. Of particular interest are the following:

- The extent to which the Offeror presents a technical team, with outstanding technical credentials, whose qualifications are clearly matched to the technical and programmatic approaches presented; and

- The extent to which the Offeror identifies and describes multiple examples of substantial, successful prior work by the Offeror and the Offeror's organization that is highly relevant across the scope of the program description in the area of Advanced Question Answering.

7.8.2 Potential Contribution and Relevance to VACE R&D Program and Research Goals

When assigning one of the four adjectival ratings listed above, each reviewer will consider their assessment of the extent to which the Offeror's proposal demonstrates knowledge and understanding of the goals, objectives and technical problems elaborated by the Government for the VACE program (See Sections 2 and 4). Of particular interest is the extent to which the Offeror's proposal describes an approach that will have a clear, positive, and significant impact on the VACE program, its goals and schedule. Also important is the extent to which the Offeror's proposal technically justifies proposed research goals, objectives and directions with projected experimental results that promise to greatly accelerate the development of technical capabilities for the VACE program. Proposals that provide the Government with sufficient rights to do a competitive reprocurement will be evaluated more favorably than proposals with restrictive rights.

7.8.3 Cost and Schedule Realism

When assigning one of the four adjectival ratings listed above, each reviewer will assess the extent to which the Offeror's proposal balances the relationship of the proposed costs to the proposed technical and scientific approach, staffing and mode of operation plus the relationship of the proposed technical tasks, milestones and schedule to the anticipated / perceived difficulty and innovativeness of the proposed technical and scientific approach.

7.9 Award Selection and Notification

As soon as the proposal evaluation is complete, the Offeror will be notified if their proposal is recommended for award and funded, recommended for award and not funded or not recommended for award. Proposals that are not recommended for award will be destroyed in accordance with normal NBC procedures.

Awards under this BAA will be made to responsible Offerors on the basis of the evaluation criteria above and a BEST VALUE approach to the Government. Not all proposals deemed selectable will be funded. The Government reserves the right to select for award all, some, or none of the proposals received and to fund incrementally any award instrument. The Government also reserves the right to fund all or any part of a proposal evaluated as eligible for award. Awards are subject to the availability of Government funds and will be made in the form of a contract.

If additional funds become available at a later date the Government reserves the right to fund any selectable proposal up to 12-months after the proposal submission date. Such funding of selectable proposals may require the Government to request specific

modifications to the technical proposal and to enter into negotiations to resolve any issues and related adjustments to the cost proposals.

APPENDIX A - Video Data Types

The VACE program anticipates that the majority of the video corpora that will be collected, developed and labeled for use by the research investigations of Phase III will be in MPEG 1 and MPEG 2 formats. In selected cases the video corpora for Phase III may be in VHS format. At the beginning of Phase III, it may be possible to shift the format to MPEG 4 or H.264 depending on the acceptance of these video standards, and the availability of relevant tools to support research.

Recent research in video has overwhelmingly focused on relatively short video clips or on still frames extracted from video clips. The intent of Phase III of VACE is have emphasis on the processing, analysis, content extraction and metadata annotation of significantly longer video segments. The collection of video data that will be used in VACE Phase III will reflect this emphasis.

In addition, the content of the video used in Phase III of VACE will have significant variation across a number of different data dimensions. These include each of the types of video briefly described in the following subsections. Each type of video may be recorded under a wide variety of conditions (e.g. lighting, times of day, weather and climatic conditions).

News Broadcast Video

News broadcast data is available from a number of sources via the services of the DNI Open Source Center and must be used in VACE under fair use provisions of the copyright act. In general, approaches that do not rely on a single audio language have priority, as foreign language broadcasts are preferred. In some cases, English language broadcasts may be available to aid in the comparative evaluation of algorithms on foreign language source but only limited translation and transcript services will be available from FBIS. Broadcasts for VACE Phase-III will be from multiple sources in Arabic, Chinese, German, and Spanish. Generally the VACE collections will be one-hour to one-half-hour broadcasts for each source over a two-week period (10-20 hours). Data will be in MPEG-2 format but it will likely exhibit artifacts of upstream compression and in some cases degradation from digitization from earlier analog (i.e. VHS) media. VACE contractors will have access to DNI Open Source services via the Internet.

Meeting/Conference Video

Meeting/Conference video covers indoor meetings in a variety of venues, formats, and topics. The venues may range from small informal settings to large conference halls seating hundreds of people. The format of the meetings will range from informal gatherings, structured educational presentations, to press conferences. These meetings may involve group discussions, visual presentations, and question/answer sessions. The topic of the meetings will range from technical presentations to group planning exercises. A major source of these video clips is expected to be the meeting room collections from NIST and LDC. There may be multiple cameras for each scene,

but the cameras will be stationary. Cameras will have the capability to pan, tilt and zoom. Audio will be available for all video, but conversation and text may be in a foreign language.

Surveillance Video

Surveillance video represents the video recorded by fixed cameras located either on the interior or exterior of buildings or facilities, at entrance/exit points, and at locations having a high security interest. Cameras may be able to rotate, pan, and zoom. An area may be under surveillance of a single camera or of multiple cameras. In the latter case the camera's field of view may or may not overlap. The surveillance camera(s) may be unattended or may be under the direct control of an operator located in a centralized control room. Typically surveillance video does not have an audio track. When an audio track is present, it may simply contain a human surveillance operator commenting on the activity being recorded.

UAV Video

UAV or Reconnaissance motion imagery data represent video recorded from airborne platforms. These video cameras may have video stabilization and may have auto-focus. The video may be shot from a wide variety of downward look angles. One common characteristic of UAV video is the reconnaissance of a general geographic area or of a specific facility (e.g. office buildings, factories, ports, or airfields). Another common characteristic of UAV video is to focus on an event or activity or to track the motion of people, vehicles or other moving objects. As in the case of surveillance video, the camera may or may not be under the direct control of the user of the data. The camera may be able to rotate, pan, tilt, and zoom independently from the motion of the airborne platform. The video's audio track, if it exists, may or may not be related to the scene being simultaneously recorded.

Current UAV video includes both human pedestrians and vehicles including tanks, missile launchers, trucks, and utility vehicles. This collected data includes:

- Time series of vehicle positions as measured by GPS receivers;
- Video from an EO camera with zoom lens mounted on a pan/tilt platform on a 300-foot tower;
- Video and metadata from EO, IR and Multi-Spectral cameras alternately flown on a Pointer UAV at altitudes of 0 to 300 ft AGL;
- Video and metadata from EO and IR cameras simultaneously flown at altitudes of 1500 to 5000 ft AGL;
- SAR data acquired at altitudes of 8500 to 12,000 ft AGL.

UAV video will typically contain one or more moving objects. Moving objects that occupy 5-35% of the field –of-view are recognized as significant.

One possible application would be the collection of video following a military action/strike or of a terrorist incident for the purposes of performing an assessment of the resulting impact, damage, and/or destruction (i.e. destruction of

buildings/vehicles/facilities, craters, broken glass and twisted metal, burning debris, casualties, and other weapon/explosive impact).

The Government will provide, upon request, UAV Motion Imagery Data as GFI to Offerors whose proposals are selected for funding under VACE Phase III. Some data from Micro UAV platforms will also be available. Data will be provided as MPEG-2 transport stream with KLV metadata. The video will be compressed using MPEG-2 standards.

Ground Reconnaissance Video

Ground Reconnaissance video clips represent recordings from hand held cameras or by cameras mounted on a moving car/vehicle. In the first case, the individual holding the camera may be stationary, walking, or riding in a vehicle. The video may or may not have stabilization or auto-focus. The recorded scenes may be indoor or outdoor and the subject of the video may be a geographic area, a facility/building(s), an event/activity, or may be a tracked person or vehicle. The distinguishing features of this video type are that the video is recorded in a more ad hoc, informal manner; typically using a hand-held video camera operated by an amateur operator. These ground reconnaissance video clips typically have a specific objective in mind (e.g. to produce a geographic or facility overview; to track a given vehicle or person; to perform a route reconnaissance; and/or to opportunistically record an unfolding event).

APPENDIX B – Exploratory Evaluation Task

Video Extraction Exploratory Task Proposals

You should submit a proposal for each exploratory task you'd like to pursue using the template below. Please submit one filled template per task. Proposals will be evaluated on the following criteria.

Criteria

Proposed task types should satisfy the following criteria:

- relevant to the need for the extraction of information from video which is useful for intelligence analysis
- able to be automatically produced with some level of accuracy by near-future systems
- able to be consistently and efficiently automatic- or hand-annotated to create a gold standard for evaluation

Two forms of exploratory tasks will be considered:

1. Component-level detection/tracking tasks along the lines of the current Core Evaluation tasks.
2. Higher-level tasks with an event or semantic analysis focus. Video event-level tasks will be most strongly considered.

Data

Below are a list of question that must be addressed in your proposal for any task you propose.

- What is the minimum optimal amount of data required?
- What data domain is needed?
- If the Proposer plans to collect data themselves, what are the:
    - Costs;
    - Distribution limitations;
    - Collection Time requirements and;
    - Impact to research progression?
- If for some reason the data collection listing is delayed what is the impact to the research?

Exploratory Task Proposal Template

Please try to address the new task using the following template as completely as possible. However, if you cannot fill in all the fields, please fill in what you can, based on your current understanding of the task you are proposing.

Complete the following template for each proposed exploratory task type:

1. Champion: *(your name and email address)*

2. Task name: *(What would you call this task?)*

3. Task Description: *(What is the task in 1 or 2 sentences?)*

4. Task data type attributes and allowable attribute values: *(How would you describe the task ontology?)*

5. Justification: *(How is this task pertinent to the above criteria?)*

6. Rules for annotation: (*How would you annotate this?)*

7. Required core technologies: *(What component technologies will this depend on?)*

8. Suggested Metric(s): *(How might the task be evaluated?)*

9. Applicable domains: *(meetings, broadcast news, surveillance, ground recon, UAV, etc.)*

10. Required resources: *(What source data and annotated development data would be needed?)*

Other notes and references to existing or contributing work:

Example filled template

1. Champion: John Doe, JohnDoe@nowhere.org

2. Task name: Vehicle tailgating

3. Task description: Detect event when one vehicle is following another vehicle without authorization.

4. Task data type attributes and allowable attribute values:
    - Object level attributes
        1. “name” (required) vehicle
        2. “type/value” (required): String (for example: Car, SUV, Motorbike)

3. “comment” (allowable): Something out of normal if any

- Sub-event level attributes: (must appear in sequence to be determined as a tailgating event)
  1. Approach {A, B}, B approaches A
  2. Access {A}, A gets access to secure area
  3. Enter {A}, A enters secure area
  4. ¬Access {B}, B does not get access to secure area
  5. Enter {B}, B enters secure area

5. Justification: The automatic detection of this event will be useful in maintaining both physical security around vehicles and in identifying when vehicles are being surreptitiously followed.

6. Rules for annotation:
   a. <object type=”OBJECT” id=”CAR1”>
      <property name=”vehicle” value=”CAR” </property>
      </object>
   b. <object type=”OBJECT” id=”SUV2”>
      <property name=”vehicle” value=”SUV” </property>
      </object>
   c. Sub-event level: This metadata type is used to identity vehicle tailgating events: when one vehicle is closely following another vehicle without authorization.
      1. <event type=”APPROACH” id=”EVENT_S1”>
         <begin unit=”frames” START_FRAME_S1 </begin>
         <end unit=”frames” END_FRAME_S1 </end>
         <argument argNum=”1” value=”CAR1” />
         <argument argNum=”2” value=”SUV2” </argument>
         <\event>
      2. <event type=”ACCESS” id=”EVENT_S2”>
         <begin unit=”frames” START_FRAME_S2 </begin>
         <end unit=”frames” END_FRAME_S2 </end>
         <argument argNum=”1” value=”CAR1” />
         </event>
      3. <event type=”ENTER” id=”EVENT_S3”>
         <begin unit=”frames” START_FRAME_S3 </begin>
         <end unit=”frames” END_FRAME_S3 </end>
         <argument argNum=”1” value=”CAR1” />
         </event>
      4. <event type=”NO_ACCESS” id=”EVENT_S4”>
         <begin unit=”frames” START_FRAME_S4 </begin>
         <end unit=”frames” END_FRAME_S4 </end>
         <argument argNum=”1” value=”SUV2” />
         </event>
      5. <event type=”ENTER” id=”EVENT_S5”>
         <begin unit=”frames” START_FRAME_S5 </begin>

<end unit="frames" END_FRAME_S5 </end>
<argument argNum="1" value="SUV2" />
</event>

d. <vehicle_tailgate> EVENT_S1, EVENT_S2, EVENT_S3, EVENT_S4, EVENT_S5 </vehicle_tailgate>

7. Required core technologies: Vehicle detection and tracking.

8. Suggested metric: Vehicle tailgate event error (Primary): Proportional to identifying each sub-event level. i. e,

$$\text{Score} = \frac{Max\ (\text{Length}\ \{\text{Correct sub - event sequence}\})}{\text{Total sequence length describing main event}}$$

From the measure, it is easy to see that if all the sub-events are identified in sequence, then the parsing will come out identified as a vehicle tailgating event with a score of 1.

For example: say three sub-events are identified successfully and the rest not identified or out of sequence, then in this case, the score will be (3/5).

9. Applicable domain: Surveillance and Aerial reconnaissance (These set of events can potentially be modified to ***people*** tailgating secure areas)

10. Required resources: We would require a corpus with approximately 50 tailgating instances per domain. We suggest beginning with ground surveillance. We will annotate the data with ground truth ourselves.

11. Other notes and references to existing or contributing work:
Jane Smith has performed some initial work in this area. See:
Smith, J. (2003), Initial experiments in automatic vehicle detection, Proc. Video Processing Workshop (pp. 35-40), Las Vegas, NV.

APPENDIX C – Evaluation Protocol

See Website for external document:

http://www.nbc.gov/fort_h/vaceiii/

APPENDIX D - TRL Definitions

TRL Levels

Tailored to DTO Technologies

Level	Definitions	Entry Condition	Contractor Activity
1	Basic principles observed and reported	Some peer review of ideas	Reporting on basic idea
2	Technology concept and/or application formulated	Target applications are proposed	Speculative work; invention
3	Analytical and experimental critical function and/or characteristic proof of concept	Algorithms run in contractor labs and basic testing is possible (internal, some external may be possible)	Doing Analytical studies with weakly integrated components
4	Component/breadboard validation in lab	Proof of concept exists; test plans exist; external testing is possible	Low fidelity integration of components
5	Component/breadboard validation in relevant environment	Integrated system functions outside contractor lab; some TRA tests completed	Working with realistic situations
6	System/subsystem model or prototype in demonstration in relevant environment	IC/DoD users identified; target environment defined; simulated testing possible	Demonstrating engineering (software qualities) feasibility
7	System prototype demo in operational environment	Test lab trials in simulated environment completed; installed in operational environment	Completing the product
8	Actual system completed and qualified through test and demonstration	Product completed; Test lab trial completed successfully	Releasing the product; Repairing minor bugs
9	Actual system proven through successful mission operations	Proven value-added in an operational environment	Repairing minor bugs; noting proven operational results

APPENDIX E – PART I

Video Analysis and Content Extraction (VACE) Program - Phase III

BROAD AGENCY ANNOUNCEMENT (BAA)

PROPOSAL COVER SHEETS

For

VOLUME 1 - Technical / Management Details

And

VOLUME 2 – Additional Reference Information

FOR Govt Use Only [Proposal Number] _____ _____

Video Analysis and Content Extraction (VACE) Program - Phase III

Broad Agency Announcement (BAA)

(Choose One)

VOLUME 1 - Technical / Management Details
or
VOLUME 2 – Additional Reference Information

Organization / Company	
CAGE Code	
DUNS / CEC Number	
TIN Number	
Type of Business	
Proposal Title	
System Design Perspective Category (Check ONLY ONE Box) (See Section 3)	❑ **1. Tier 1** ❑ **A. Metadata Content Extraction** ❑ **B. Intelligent Content Services** ❑ **C. Enabling Technologies** ❑ **D. Other Technologies** ❑ **2. Tier 2 - End-to-End system** ❑ **A. System Level** ❑ **B. Large component**

Team Members / Type of Business	
Principal Investigator(s) Name(s)	
Mail Address	
Phone Number	
Fax Number	
E-mail Address	
Administrative Contact Name	
Mail Address	
Phone Number	
Fax Number	
E-mail Address	

Proposal Duration	
Cost - Year 1	**$**
Cost - Year 2	**$**
Cost - Year 3	**$**
Total Cost	**$**

APPENDIX E – PART II

Video Analysis and Content Extraction (VACE) Program - Phase III

Broad Agency Announcement (BAA)

PROPOSAL PRICING SHEET

FOR Govt Use Only [Proposal Number] _____ _____

Video Analysis and Content Extraction (VACE) Program - Phase III

Broad Agency Announcement (BAA) BAA 06-01-MT

VOLUME 3 – Cost Information

1. Company / Agency Information:

_____
(Company / Agency Name)

_____
(First Line of Address)

_____
(Street Address)

_____ _____ _____
(City) (State) (Zip Code)

2. Company / Agency Point of Contact Information:

_____ _____
(POC Name) (POC Title)

_____ _____
(POC Telephone and FAX Nos. (Include Area Code)) (POC e-mail)

3. Type Of Contract (Check One):

____ FFP ____ CPFF ____ CPAF

____ FPI ____ CPIF ____ Other (Specify)

4. Proposed Cost (A + B = C):

	4.a. Cost	4.b. Profit / Fee	4.c. Total
Year 1	_____	_____	_____
Year2	_____	_____	_____

Total Proposed _____

BAA 06-01-MT - VOLUME 3 - Cost Information (CONTINUED)

5. Performance:

5.a. Place (1) _____ 5.b. Period (1) _____
(2) _____ (2) _____

6. Line Item Costs (List and reference the identification, quantity and total price proposed for each contract line item. A line item cost breakdown supporting this recap is required unless otherwise specified by the Contracting Officer. Continue on reverse, and then on plain paper, if necessary. Use same headings.)

6.a. Line No.	6.b. Identification	6.c. Quantity	6.d. Price	6.e. Prop. Pg. No.

7. Provide the Following (If available):

_____ _____
(Name of Contract Administration Office) (Name of Audit Office)

_____ _____ _____ _____ _____ _____
(City) (State) (Zip Code) (City) (State) (Zip Code)

_____ _____
(Telephone (Include Area Code)) (Telephone (Include Area Code))

8. Will you require the use of any Government property in the performance of this work? ____Yes No

9. Do you require Government contract financing to perform this proposed contract? ____Yes ___No
Type of financing (Check One) ____Advanced Payments ____Progress Payments ____Guaranteed Loans

10. Have you been awarded any contracts or subcontracts for the same or similar items within the past 3 years?
____Yes ____No (If "Yes," identify items(s), customer(s) and contract number(s) on reverse of form.)

11. Is this proposal consistent with your established estimating and accounting practices and procedures and FAR Part 31, Cost Principles? ____Yes ____No (If "No," explain on reverse of form.)

12. Cost Accounting Standards Board (CASB) Data (Public Law 91-379 as amended and FAR Part 30)

12.a. Will this contract action be subject to CASB regulations? ___Yes ____ No
(If "No," explain on reverse of form.)

12.b. Have you submitted a CASB disclosure statement (CASB DS-1 or 2)? ____Yes ___No
(If "yes," specify in proposal the office to which submitted and if determined to be accurate.)

12.c. Have you been notified that you are or may be in compliance with your disclosure statement or cost accounting standards? ____Yes ____No (If "Yes," explain in proposal.)

12.d. Is any aspect of this proposal inconsistent with your disclosed practices or applicable cost accounting
standards? ____Yes ____No (If "Yes," explain in proposal.)

This proposal is submitted in response to BAA 06-01-MT and reflects our estimates and/or actual costs as of this date and conforms to the instructions in FAR15.804-6(b)(1), and Table 15-2. By submitting this proposal, the Offeror, if selected for negotiation, grants the contracting officer and authorized representatives(s) the right to examine, at any time before award, those records which include books, documents, accounting procedures and practices, and other data regardless of type and regardless of whether such items are in written form, in the form of computer data, or whether such supporting information is specifically referenced or included in the proposal as the basis for pricing, that will permit an adequate evaluation of the proposed price.

13. Name (Typed)	**14. Title**	**15. Company / Agency Name**

16. Signature	**Date**