



I A R P A  
BE THE FUTURE

## **ALADDIN Proposers Day IARPA-BAA-10-01**

---

**Automated Low-level Analysis and Description of Diverse Intelligence Video**

Dan Aldridge

Program Manager

Intelligence Advanced Research Projects Activity (IARPA)

301-226-9100

email: [daniel.g.aldridge@ugov.gov](mailto:daniel.g.aldridge@ugov.gov)



# Disclaimer

---

- **This presentation is provided solely for information and planning purposes.**
- **The Proposers' Day Conference does not constitute a formal solicitation for proposals or proposal abstracts.**
- **Nothing said at Proposers' Day changes the requirements set forth in a BAA.**
- **Any conflict between what is said at Proposers' Day and what is in a BAA will be resolved in favor of the BAA.**



# No White Papers

---

- **No White Papers will be requested for ALADDIN**
- **Proposals will be due approximately 45 days after BAA is published**
- **Take advantage of this time to start developing your ideas**



# Outline

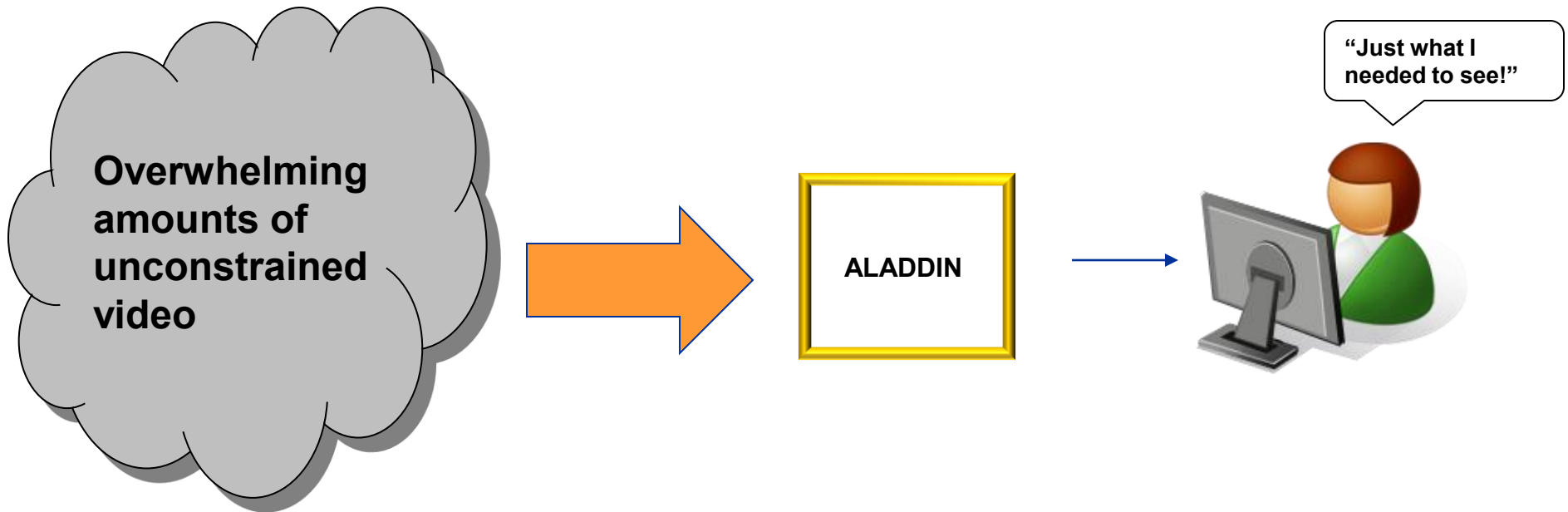
---

- **Program Overview**
- **Program Phases, Metrics, and Milestones**
- **Award Information**
- **Eligibility Information**
- **Application Review Information**

**Question periods sprinkled throughout**



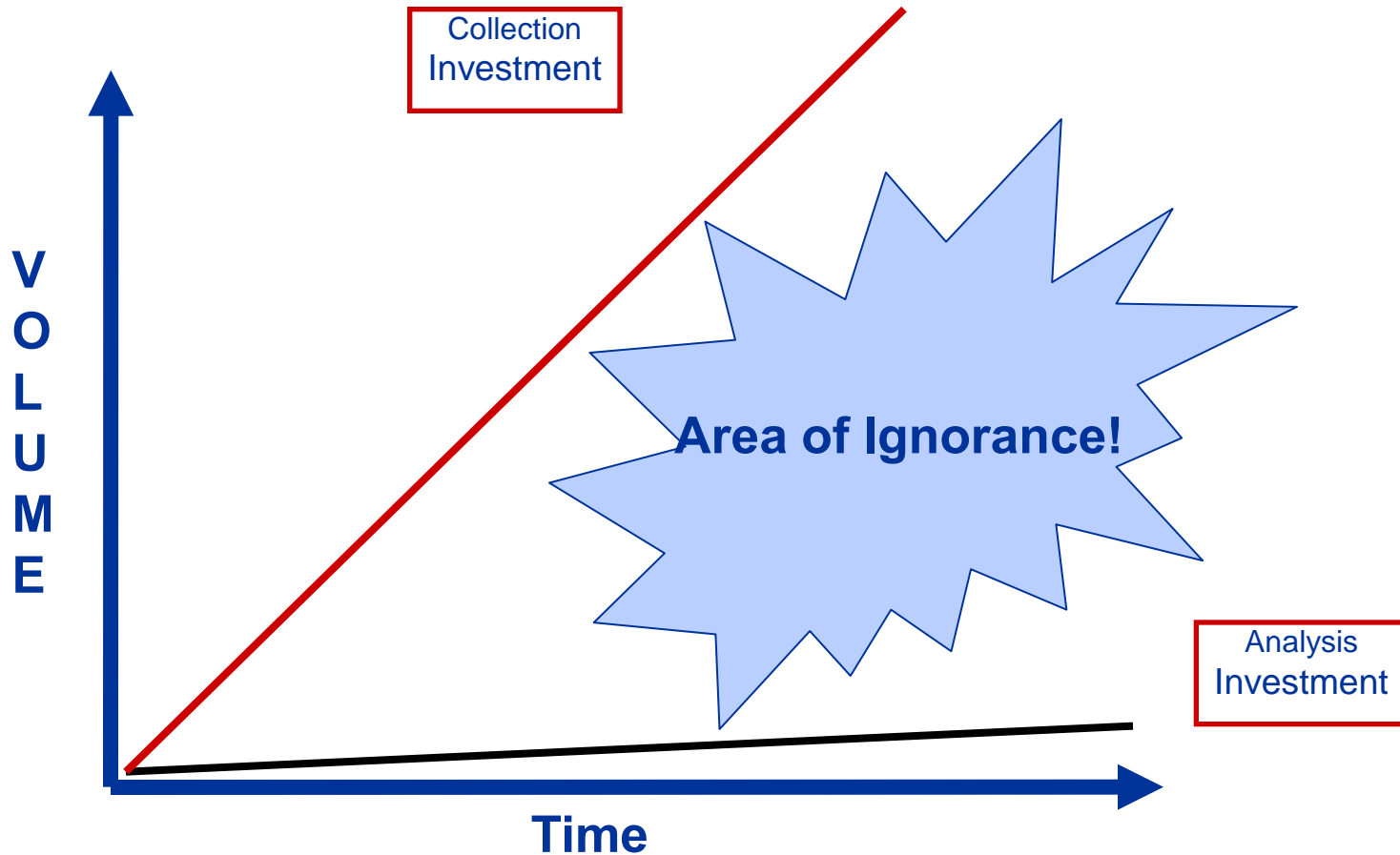
# ALADDIN Program Overview





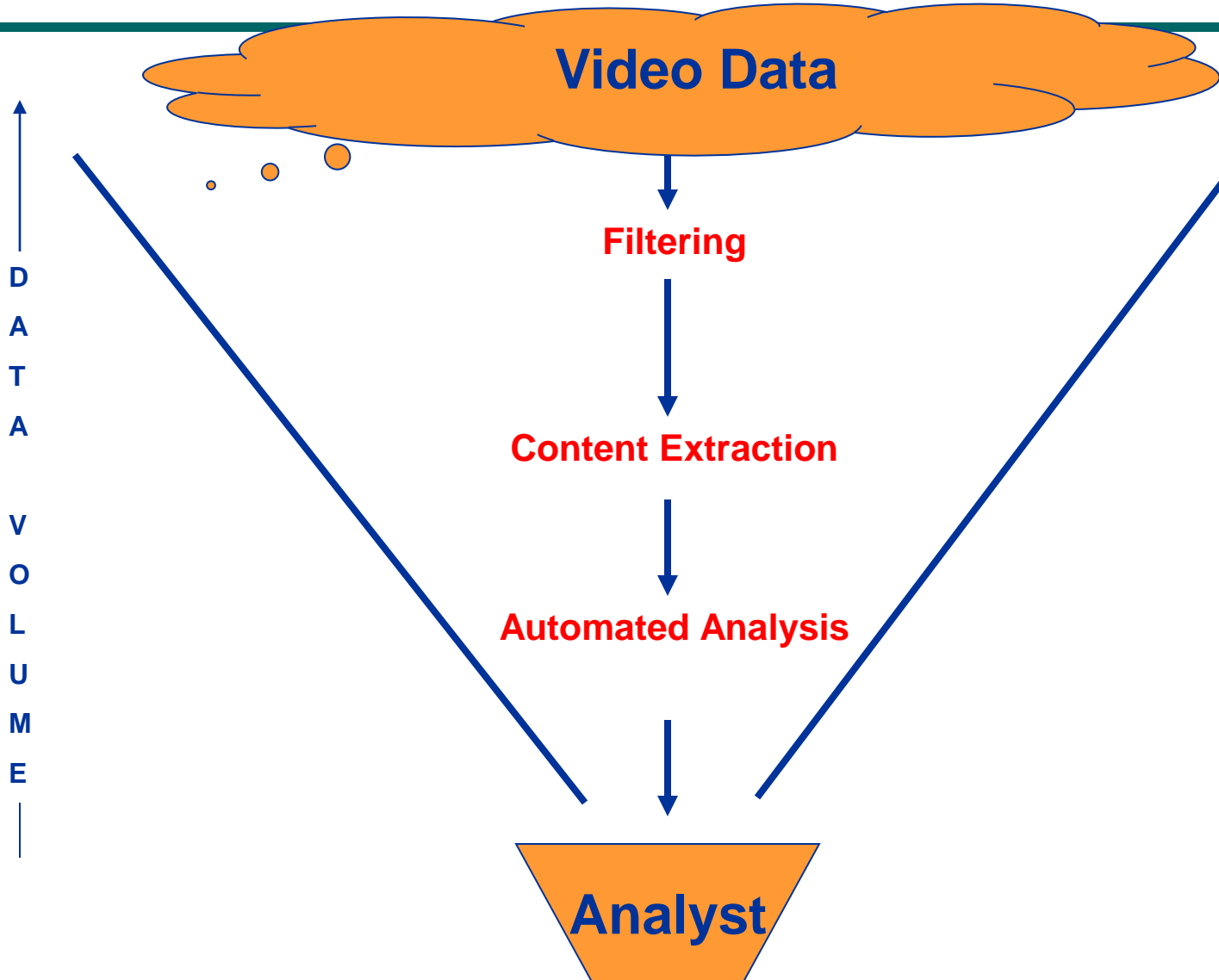
# Video Analysis Problem

**We don't know what we don't know!**





# The Video Data Funnel





## ALADDIN Program Goal

---

**The overarching goal of ALADDIN is to allow Intelligence Community analysts to quickly find events of interest from unconstrained video (videos for which we do not control the sensor).**





# Program Impact Metrics

---

- **40X throughput speed-up per analyst**
- **Analysts working with an ALADDIN-based system find 80+% of the Events of Interest in very large video collections**
- **100% of videos will be “processed”**



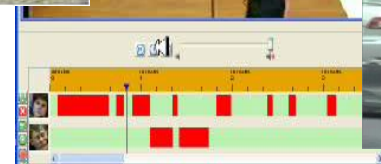
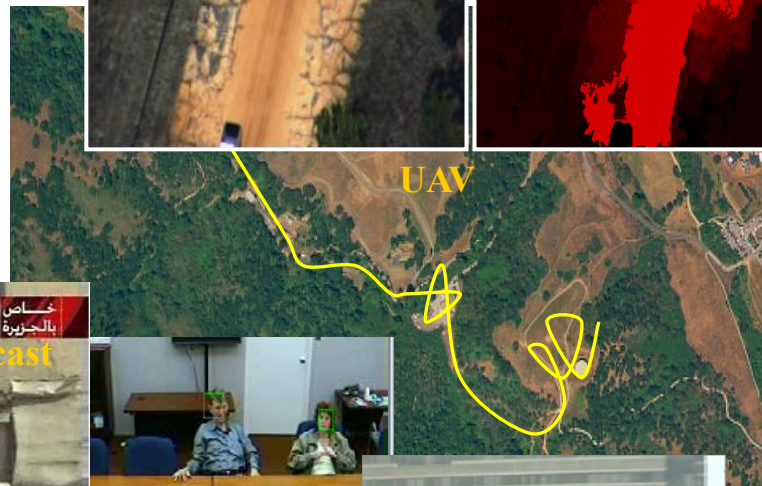
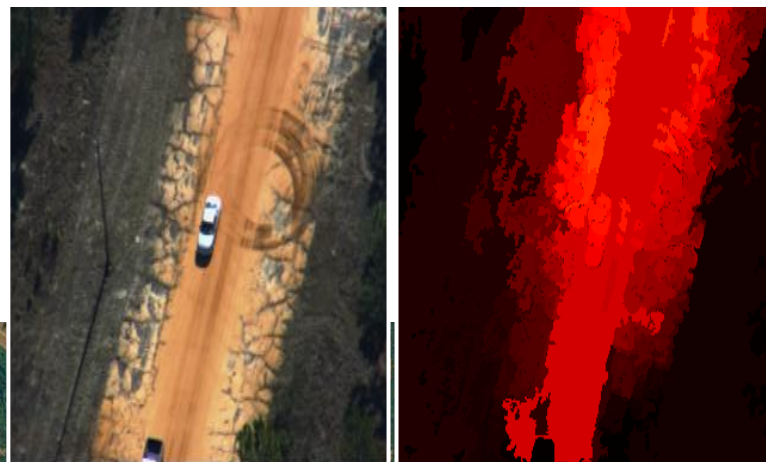
# State of the Practice - Domain

- Broadcast
- Surveillance (and some Ground Reconnaissance)
- Video from Unmanned Aerial Vehicles (UAV's)
- Meetings, Interviews, etc.

All either constrained or highly formatted.



يتكون  
ك  
ه  
و

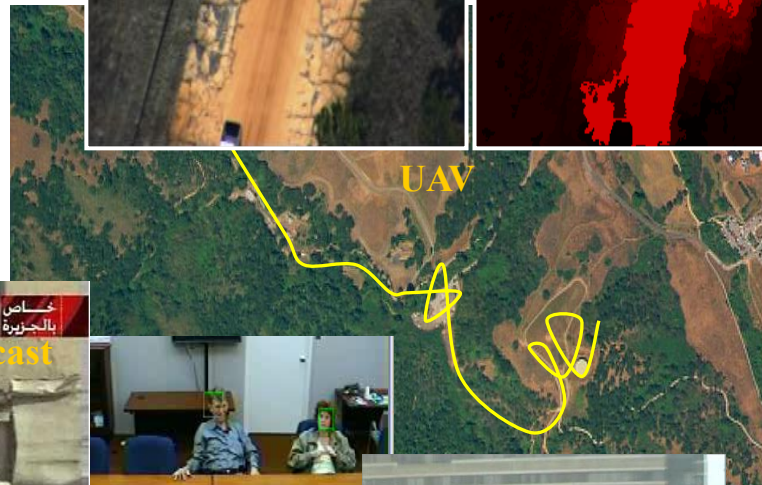
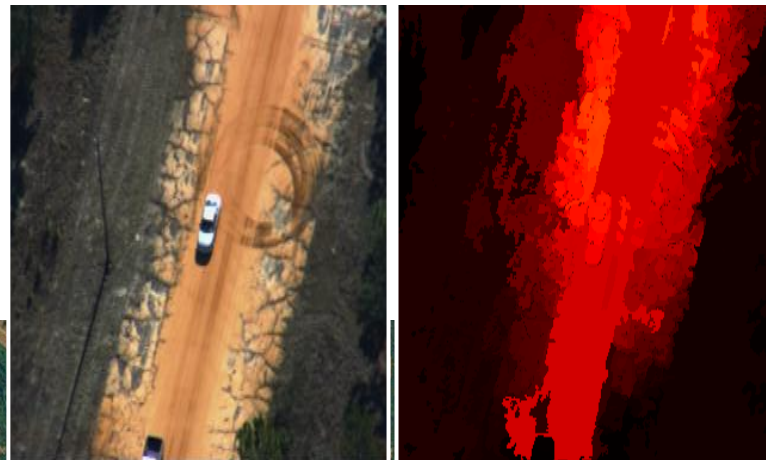




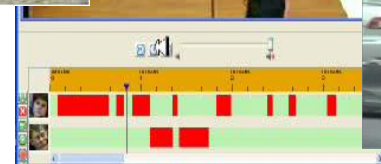
# State of the Practice - Detection

- Detection of People, Vehicles, Text, Various other objects
- Detection of certain actions/motions

Generally, all this detection has been within the constrained data set.



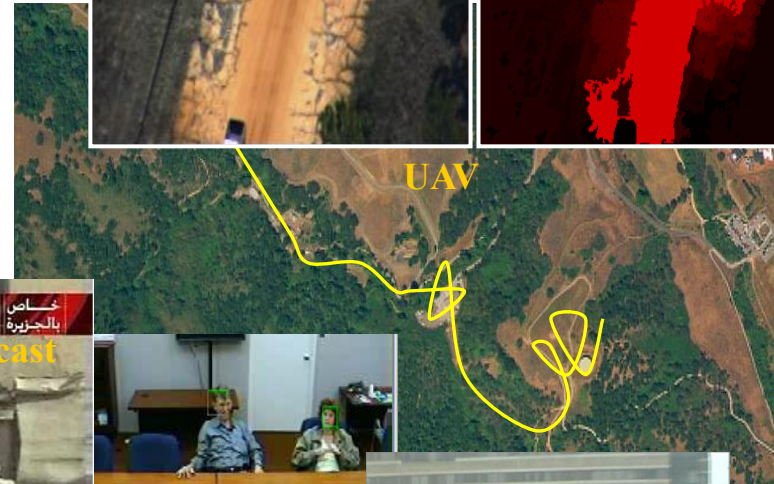
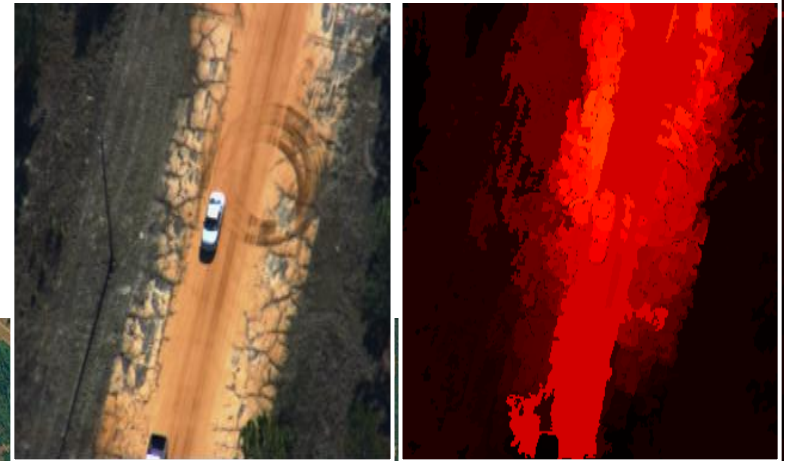
يتكون  
من





# State of the Practice – Content Based Retrieval

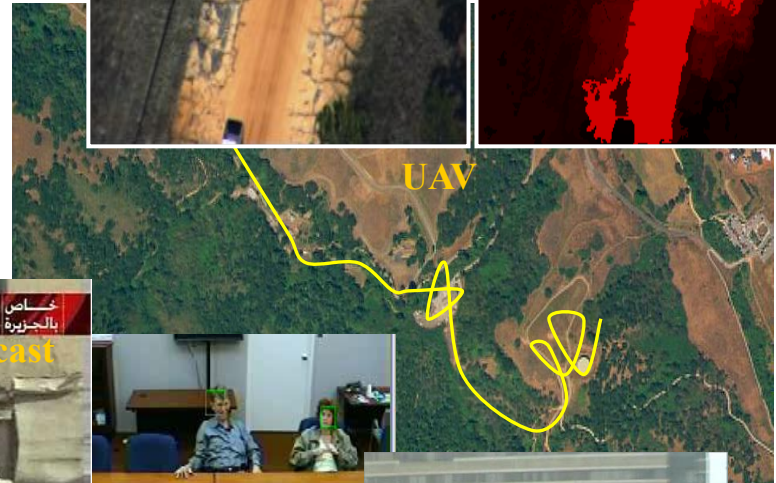
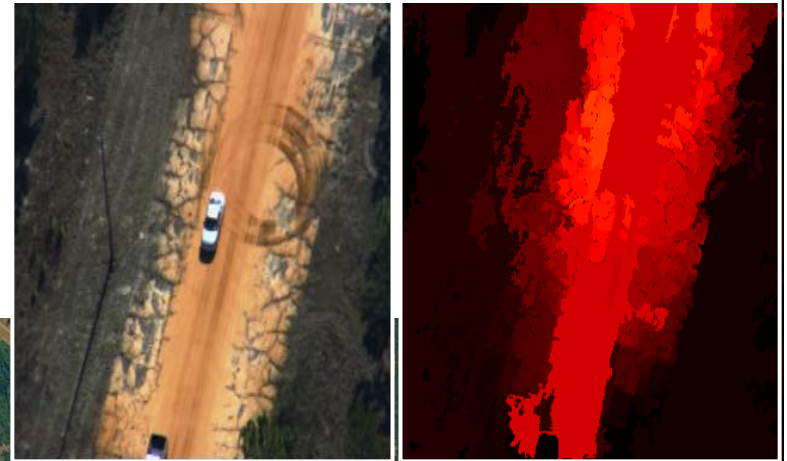
- Detection of general independent features that generate a “bag of words” searchable lexicon from video and audio
- Relevance-based retrieval ranked on the detected features





# State of the Practice – Events and Unconstrained Data

- Detection of objects within the unconstrained domain
- Detection of events within the constrained domain



يتكون  
من



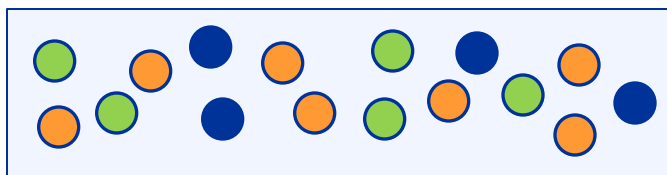


# ALADDIN Vision

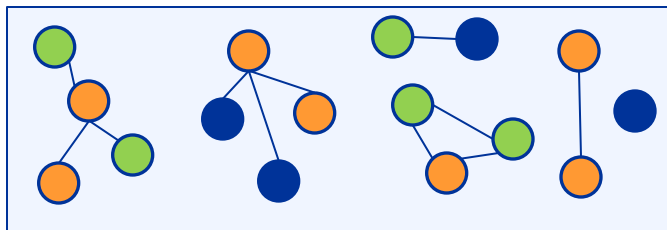
## Filtering



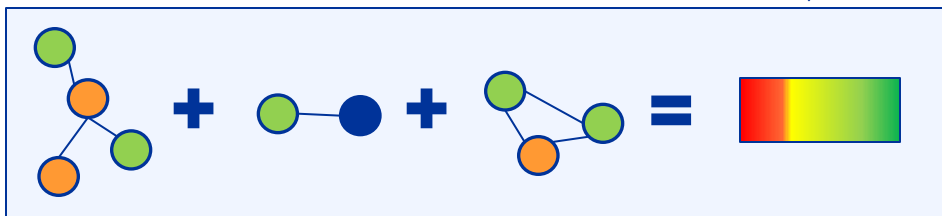
## Metadata Generation



## Content Description



## Detection of Events of Interest



## Recounting of Events of Interest





# ALADDIN Key Components

---

## Filtering: ~10%

- Create non-content based filters to identify video clips that have no analytic value and can be eliminated from content analysis. For example, the file name, the file length, the extension type, etc., may all strongly suggest that the file will not contain an event of interest and does not require further processing. (surface signal characteristics)

## Metadata Generation: ~20%

- Extend current state-of-the-art techniques and build necessary new ones to analyze the salient content within these video files and generate metadata to support accurate detection of the specified events. Types of metadata we might use might include certain objects (such as bicycle, car, building, road); actions (such as running, throwing, setting down); and sounds (applause, explosions, screams, etc.).

## Content Description: ~30%

- Create a structural representation of the metadata to “describe” the extracted content which is enriched by relational/hierarchical/temporal relationships. This will be within short (probably 2-3 minute), individual, video files, not across a corpus of video data.

## Detection of Events of Interest: ~30%

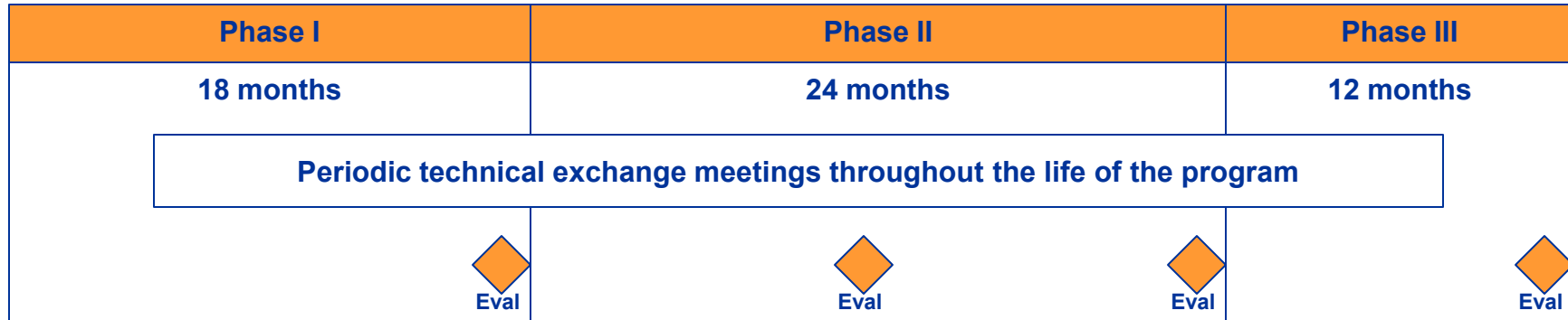
- Utilize the content description to accurately detect the occurrence of the events

## Recounting of Events of Interest: ~10%

- Generate an account of the event in human-readable text providing key evidential detail. The plan is to let each research team create their own form of recounting, seeking the best way to help the human understand what event of interest is contained within the video file.



# Program Structure



- Develop prototypes of all ALADDIN Key Components
- Tested against predefined events
- Capability focused

- Continued R&D on all ALADDIN Key Components
- Annual evaluations against predefined and novel events
- Capability and Performance focused

- Prepare ALADDIN Key Components for final evaluation
- Tested against predefined and novel events
- Capability, Performance and Integration focused

All research data will be unclassified

IARPA will provide substantial research data sets to ALADDIN performers





# Research Deliverables

---

- **Core metadata generation and analysis algorithms developed in the course of research**
- **Content Description Framework**
- **Event detection algorithms**
- **Event recounting algorithms**
- **Documented process for development of novel event detection algorithms**
- **Evaluation runs on data and system output as specified, including content descriptions, event hypotheses, and event accounts.**

**Concentrate on development of advanced automatic capabilities with regard to detection, analysis and analytic integration with the goal of accurate and robust event detection.**

**Produce a set of fully-integrated algorithms and frameworks. Note that performers are only expected to build minimal-effort integration prototypes that demonstrate potential utility.**



## Scope

---

**WILL support integration of the following capabilities, but not core research:**

- **Core metadata generation capabilities unless these involve novel breakthroughs critical to program**
- **Person/face/vehicle detection and tracking**
- **Speech-to-text transcription and speaker recognition/speaker ID**
- **Machine Translation**



## Scope, Continued

---

**What is the program NOT trying to do?**

- **Develop a methodology to identify specific people in video**
- **Develop a methodology to identify specific locations in videos**
- **Detect events in the “constrained” domains, i.e. ...**
  - **broadcast**
  - **surveillance**
  - **aerial/UAV**
- **Speech Processing and Analysis**



---

# Questions?



---

# ALADDIN Metrics



# Program Impact Metrics

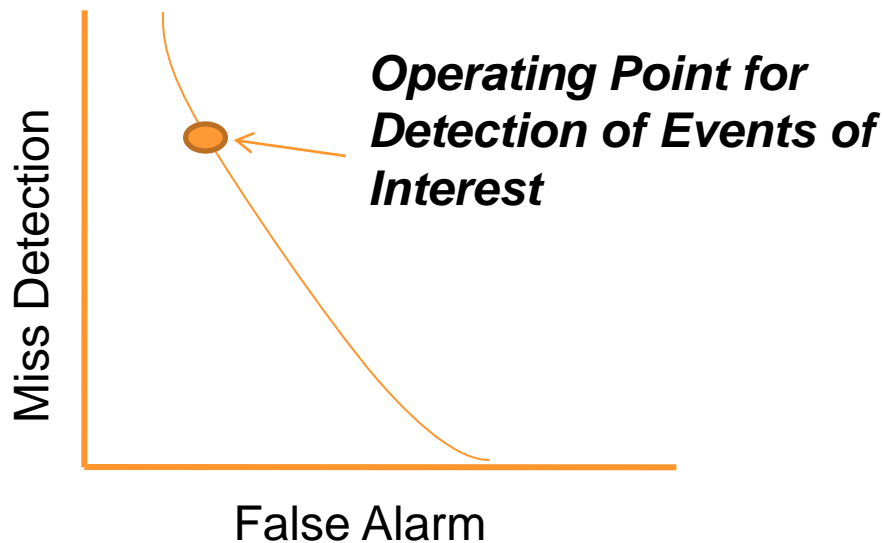
---

- **40X throughput speed-up per analyst**
- **Analysts working with an ALADDIN-based system find 80+% of the Events of Interest in very large video collections**
- **100% of videos will be “processed”**



# ALADDIN Metrics: Detection of Events of Interest

Decision Error Tradeoff (DET) Curve



**Specific Phase targets are to be determined; they will be in the BAA. We estimate they will follow a Moore's Law curve in reduction of error rate.**

- **Systems will be required to emit probabilities for each specified event of interest against each video file in the test collection.**
- **These outputs will be used both to generate diagnostic DET curves and operating point scores.**
- **The operating point scores will be the primary metrics.**
- **The Operating Point for Detection of Events of Interest is intended to support human analytic needs**
  - *Analysts prefer a low false alarm rate*



# ALADDIN Metrics: Recounting of Events of Interest

---

The test collection for Event Recounting will be a proper subset of the test collection for Detection of Events of Interest.

Perform human assessment of the event accounts as follows:

1. Is the account interpretable?
2. Based solely on the evidence in the account, did the event detector make the correct decision?
3. Utilizing both the account and source multimedia file, assess the utility of the accounts
  - Assess completeness, accuracy, clarity, conciseness, and other important dimensions on a Likert scale and report diagnostic analysis of error





---

# Questions?



---

# **ALADDIN**

## **Award Information**



# Award Plan

---

## **4.5-year Program starting 3Q FY2010**

- **Phase I – Base Period -- 18 months**
- **Phase II – Option 1 and 2 Period -- 12 months each**
- **Phase III – Option 3 Period -- 12 months**

**Criteria for advancing to next phase: sufficient progress in achieving prior phase metrics**

**Multiple awards anticipated, depending upon**

- **quality of the proposals received**
- **availability of funds**



# Teaming

---

**Because of the complexity of the program and the diversity of the research, the ALADDIN solicitation will encourage a teaming approach to solving the problem. Consider the following objectives when forming teams:**

- **Completeness**
  - **Teams should not lack any capability necessary for success. Do not rely upon results from the community at large or other enabling technology to be developed elsewhere**
- **Integrated Management**
  - **Clear, strong management, single point of contact**
  - **No loose confederations**
  - **Each team member should contribute significantly to the program research goals. Explain why each team member is important**

**Remember, you may be very accomplished, but can you do it all?**



---

# **ALADDIN**

## **Eligibility Information**



## ALADDIN Eligibility Information

---

- **All responsible sources capable of satisfying the ALADDIN program goals may submit a proposal or join with others in submitting proposals.**
- **Collaborative efforts/teaming among potential performers is encouraged where it enables program success.**
- **Foreign organizations and/or individuals may participate**
  - **Must comply with Non-Disclosure Agreements, Security Regulations, Export Control Laws, etc, as appropriate.**



# Ineligible Organizations

---

- **Other Government Agencies**
- **Federally Funded Research and Development Centers (FFRDCs)**
- **University Affiliated Research Centers (UARCs)**
- **Any other similar type of organization that has a special relationship with the Government, that gives them access to privileged and/or proprietary information or access to Government equipment or real property**



---

# ALADDIN

## Evaluation Criteria





# ALADDIN Evaluation Criteria

---

**Evaluation criteria in descending order of importance:**

- **Overall Scientific and Technical Merit**
- **Effectiveness of Proposed Work Plan**
- **Relevance to IARPA Mission and ALADDIN Program Goals**
- **Relevant Experience and Expertise**
- **Cost Realism**



# ALADDIN

---

**Thank you!**

**Final questions?**



# Point of Contact

---

**Dan Aldridge  
Program Manager  
IARPA, Incisive Analysis Office  
Office of the Director of National Intelligence  
Intelligence Advanced Research Projects Activity  
Washington, DC 20511**

**Phone: 301-226-9130**

**Fax: 301-226-9137**

**Electronic mail: [dni-iarpa-baa-10-01@ugov.gov](mailto:dni-iarpa-baa-10-01@ugov.gov)  
(include IARPA-BAA-10-01 in the Subject Line)**

**Website: [www.iarpa.gov](http://www.iarpa.gov)**