OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE

OPEN SOURCE INDICATORS (OSI) PROPOSERS' DAY BRIEFING

Office of Incisive Analysis

LEADING INTELLIGENCE INTEGRATION

Jason Matheny
August 3, 2011

Agenda

Time	Daily Event Schedule	Speaker
7:30 AM – 8:30 AM	Registration	
8:30 AM – 8:45 AM	Welcome Remarks	**Mr. Jason Matheny** Program Manager, IARPA
8:45 AM – 9:15 AM	IARPA Overview	**Dr. Lisa Porter** Director, IARPA
9:15 AM – 10:00 AM	OSI Overview	**Mr. Jason Matheny** Program Manager, IARPA
10:00 AM– 10:30 AM	**Break**	
10:30 AM – 11:00 AM	OSI Q&A	
11:00 AM – 11:15 AM	Contracting	**Mr. Kurtis Jones** Contracting Officer, IARPA
11:15 AM – 11:30 AM	Contracting Agent Overview	**Mr. Larry Carter** DOI Contracting Officer's Representative
11:30 AM – 12:30PM	**Lunch**	
12:30 PM – 2:00 PM	Capability Presentations (Government Representatives will not be present)	
2:00 PM – 3:30 PM	Networking and Teaming Discussions (Government Representatives will not be present)	Colony Ballroom (2nd Floor, Room 2203)

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.
- BAA supersedes anything presented or said at the Proposers' Day by IARPA.

Proposers' Day Goals

- Familiarize participants with IARPA's interest in open source indicators. *Please ask questions and provide feedback; this is your chance to alter the course of events.*
- Foster discussion of synergistic capabilities among potential program participants, i.e. foster teaming. *Take a chance: someone might have a missing piece of your puzzle.*

Schedule

- Full Proposals are due ~45 days after BAA is published.
- Once BAA is released, questions can only be submitted and answered in writing via the program website.

OSI Overview

- **OSI is a research and development program** with one Base Year and two Option Years. It **seeks to develop methods for continuous, automated analysis of publicly available data in order to anticipate and/or detect societal disruptions**, such as political crises, disease outbreaks, economic instability, resource shortages, and responses to natural disasters.

- **OSI will aim to develop methods that "anticipate the news"** by fusing early indicators of events from multiple publicly available data sources and types.

Background

- Many significant societal events are preceded and/or followed by population-level changes in communication, consumption, and movement.
- Some changes may be indirectly observable from publicly available data, such as web search queries, blogs, microblogs, internet traffic, webcams, financial markets, Wikipedia edits, and many others.
- Published research has found that some of these data sources are individually useful in the early detection of events such as disease outbreaks and macroeconomic trends.
- Little research has examined the value of combinations of data from diverse sources.

Current Research

- Analysis of news feeds, Twitter, blogs, and web search queries for detecting disease outbreaks, forecasting product sales & macroeconomic trends.

- IARPA's Aggregative Contingent Estimation (ACE), DARPA's Integrated Crisis Early Warning System (ICEWS), and the Political Instability Task Force: forecasts for pre-defined events.

- Few methods have been developed for detecting / anticipating unexpected events by fusing data of multiple types from multiple sources.

Key Technical Challenges

- **Development of models for population behavior change** in anticipation of, and in response to, events of interest.
- **Collection and processing of publicly available data** that reflect those population behavior changes. (We are looking for creative ideas.)
- **Development of data extraction techniques that focus on volume**, rather than depth, by identifying shallow features of data that correlate with events.
- **Development of multivariate time series models** robust to non-stationary, noisy data to reveal patterns that precede events.
- Use of **novel techniques to estimate causality in time series**.
- **Training of classifiers to weight combinations of time series** for generating probabilistic warnings of events.

Evaluation

- **Teams will deliver real-world warnings to IARPA**. The goal is to "anticipate the news."
  - Teams choose sensors, data, and methods.
  - Teams are rewarded for early and accurate warnings of as many newsworthy events as possible.
- **Warning delivered to IARPA** = **{Time stamp, Probability of event, Event description}**
- **Event description** = **(Population, Event-Type, Event-Time, Location)**
- Performers can elect to send additional details about events.
- The **delivery of successive, better warnings is encouraged**.
- Each warning will be scored separately.

Geographic Scope

- **OSI is an R&D program, not an operational global watchboard**.
- For research purposes, the **geographic focus will be Latin America, without the Caribbean**.
  - 21 countries, large enough to test the generalizability of performers' approaches
  - Sufficiently representative
    - Some countries with few events
    - Some countries with many events
  - Variety of publicly available data
  - Good "ground truth" reporting for training and testing

Events and Scoring

- At kickoff, Government team will provide a large list of significant events in Latin America for the prior year, for which an early warning would have been valuable.
- After kickoff, Government team will provide monthly "ground truth" – events for the last month, for which a warning would have been expected.
- Starting in Month 5, teams will deliver warnings to IARPA.
- Starting in Month 9, warnings delivered to IARPA are scored against Program milestones.

Events and Scoring

- Scoring:
  - **Lead time** = time of warning delivered to IARPA to time of earliest report (not necessarily time of event) by a major news source or other "Gold Standard" report.
  - **Probability score** = accuracy of probability assigned to event.
  - **Quality of warning** = typological match between event forecasted/detected and true event.
  - **Precision and Recall**.

- **Other assessments, qualitative and quantitative**, **will be performed by the Government team** to evaluate each team's approach (e.g. the scientific merit of the approach and the research findings). Most importantly, **approaches will also be evaluated on the utility of the warnings, as judged by potential users**.

Metrics

- **Lead Time**
  - **Days between warning and "Gold Standard" report.**
  - Warning for an event outside the 30-day window will not be scored, but should be submitted. Such events will be analyzed separately to provide additional assessment of the team's approach.
  - While successive, better warnings for the same event will be scored separately, teams will be asked to identify such successive warnings. The Government team will use this information for additional assessments of team's approach.

- **Probability Score**
  - **Quadratic score = $1 - (o-p)^2$**
  - p is the probability assigned to the warning, o is "ground truth": 1 if the event occurred, 0 if the event didn't occur within 30 days.

Metrics – Quality of Warning

- **For each warning we calculate the quality $q = \alpha_1+\alpha_2+\alpha_3+\alpha_4$**
  - α_1 ~ Population; α_2 ~ Event type; α_3~ Location; α_4 ~ Event time
  - This provides "partial credit" for partial warnings.
- **Use typology of populations, events, locations to calculate match between warning and ground truth**; e.g., for locations:
  - Typology = (Country, Province/State, City).
  - Compare warning location with true location to get (x_1, x_2, x_3), $x_i = 0$ if false, $x_i = 1$ if true.
  - Location quality = $\alpha_3 = 1/3\ x_1 + 1/3\ x_1x_2 + 1/3\ x_1x_2x_3$
    - If the warning has the wrong country, then $\alpha_3 = 0$.
    - If the warning has the country right but everything else wrong, $\alpha_3 = 1/3$.
    - If the warning has the country and the province/state right but the city wrong, $\alpha_3 = 2/3$.
    - If all is right, $\alpha_3 = 1$.
- **For the time of the event, use *1- min(|warning time – actual time|,30)/30***

Metrics - Example

- Warnings:

	Time stamp	Probability of event	Population	Event-Type	Event-Time	Location
Warning1	11/03/11	0.15	General population	Flu outbreak	11/03/11	Brazil
Warning2	11/05/11	0.30	General population	Flu outbreak	11/03/11	Paranaguá
Warning3	11/09/11	0.45	General population	Flu outbreak	11/03/11	Curitiba
Gold Standard	11/11/11	1	General population	Flu outbreak	11/03/11	Curitiba

- Quality scores (location):

Level	Gold Standard	Warning1	Warning2	Warning3
Country	Brazil	Brazil	Brazil	Brazil
Province/State	Paraná	-	Paraná	Paraná
City	Curitiba	-	Paranaguá	Curitiba
Score	**1**	**1/3**	**2/3**	**1**

- Overall scores:

Metric	Warning1	Warning2	Warning3
Lead time	8 days	6 days	2 days
Probability score	0.28	0.51	0.70
Quality score	3.33	3.67	4

Metrics

- **Precision**:
  - (Number of events identified by Government team for which performer team sent a warning to IARPA with non-zero lead time and quality) / (total number of warnings sent to IARPA by performer team)

- **Recall**:
  - (Number of events identified by Government team for which performer team sent a warning to IARPA with non-zero lead time and quality) / (total number of relevant events identified by Government team)

Warning Generation

- It is expected that the **technology developed under this effort will have no "human in the loop."**
- The teams' systems will generate warnings without the help of SMEs, either to guide the system or to filter warnings before they are sent to IARPA.
  - SMEs can help develop and train the system.
- Teams' **systems must include an audit trail for each warning**, listing relevant evidence and weights.
- Warnings that are related should be explicitly identified for additional evaluation by the Government team.
  - E.g. successive warnings for the same event, warning for mutually exclusive events.

Milestones

- Month 5 – Deliver first warnings (ungraded waypoint)
- Month 9 – Deliver first warnings (graded against milestones)
- Year-end milestones:

Metric	Month 12 (4 months of warnings)	Month 24 (12 months of warnings)	Month 36 (12 months of warnings)
Mean Lead Time	1 day	3 days	7 days
Mean Quality Score	3	3.25	3.5
Mean Probability Score	0.60	0.70	0.85
Recall	0.50	0.65	0.80
Precision	0.50	0.65	0.80

What OSI is not

- **The OSI Program will <u>not</u> fund research on**:
  - U.S. events
  - Identification or tracking of specific individuals
  - Collection mechanisms that require directed participation by individuals

What OSI is not

- **OSI is not a program on advanced NLP**
  - Off-the-shelf NLP should be sufficient for extracting features of interest
  - Deeper parsing needed for reducing error in analysis of individual signals; but OSI approach should cancel error through aggregation, focusing on population-level behavior
- **OSI is not narrowly focused on a single data source or type**
  - OSI is not about developing advanced tools for analyzing a single signal
  - OSI is about developing new methods for aggregating multiple, noisy signals indicative of significant events

Data

- Data acquisition/collection will require some resources (time and budget) by each team, and data requirements will likely overlap across teams.
- To maximize the use of resources towards technology development, IARPA may make a data investment.
  - BAA will ask that bidders list data sources required for their approach and that cost proposals include estimates of all data costs.
  - After source selection IARPA will identify core data sets -- the intersection of selected teams' data requirements, justified by the most promising technical approaches.
  - IARPA may acquire some core data sets and make them available to all selected teams.
  - Performers will adjust cost proposals to reflect IARPA data investments.

Teaming

- Because of the many challenges presented by this program, both depth and diversity will be beneficial.
  - Throughput. Consider all that you will need to do, all the ideas you will need to test.
    - Make sure you have enough people and expertise to do the job.
    - Sufficient resources to follow critical path while still exploring alternatives – risk mitigation
  - Completeness. Teams should not lack any capability necessary for success, e.g. should not rely on enabling technology to be developed elsewhere
  - Tightly knit teams.
    - Clear, strong, management, and single point of contact
    - No loose confederations
    - Each team member should be contributing significantly to the program goals. Explain why each member is important, i.e. if you didn't have them, what wouldn't get done?
    - No teaming for teaming's sake
- Remember, you may be very accomplished, but can you do it all?

Team Composition

- Given the combination of technical challenges, we anticipate teams will possess expertise in:
  - Social sciences
  - Mathematics and statistics
  - Computer science
  - Signal processing
  - Content extraction
  - Information theory
  - Software rapid prototype development

Eligibility Information

- Other Government Agencies, Federally Funded Research and Development Centers (FFRDCs), University Affiliated Research Centers (UARCs), and 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, are not eligible to submit proposals under this BAA or participate as team members under proposals submitted by eligible entities.

- Non-US organizations and individuals may be able to participate.

  – Must comply with Non-Disclosure Agreements, Security Regulations, Export Control Laws, etc, as appropriate

  – Specific guidance for non-US participation will be provided in the BAA.

Proposal Guidance

- Your proposal should include a **full discussion of the technical approach** that will be used to meet the program goals.

- Programmatic issues that should be addressed in the proposal:
  - Your **team's current technical capabilities**
  - **Key resources needed** (not currently available to your team), to include capital equipment and special expertise (teaming will likely play an essential role in providing special expertise). The risk in acquiring these key resources, and mitigation strategies, should be indicated as well.
  - A **teaming plan** along with the roles and responsibilities of each member of the research team
  - **Annual milestones** and some intermediate waypoints are set, but it is **expected that other intermediate waypoints are on the critical path** of the proposed approach will be offered.
  - A **schedule of all milestones including a clearly charted description of** the **various risk mitigation strategies** that will be undertaken to achieve program goals

Proposal Guidance (contd.)

- Download ALL materials posted to the FedBizOpps announcement (BAA, instructions, templates, etc.).
- Periodically check for amendments and other information that may be posted prior to the proposal due date.
- Read FAQs posted to the IARPA OSI Program website.
- Ensure submission requirements are followed:
  - Deadlines
  - Do not exceed page limits
  - Use all provided templates (see Appendix)
  - Include all required responses (OCI paperwork, Academic Acknowledgement letters, etc.)
  - No unnecessarily elaborate brochures or marketing material
  - Failure to follow the submission procedures may result in the submission not being evaluated

Proposal Evaluation Criteria

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

Additional Information

- dni-iarpa-baa-11-11@ugov.gov for additional questions
- OSI BAA will be posted on the FedBizOpps website (www.fedbizopps.gov)
- Q&As will appear after the BAA on the IARPA OSI Program website (http://www.iarpa.gov/solicitations_osi.html)