NOTE: Due to valuable feedback as a result of Proposers’ Day, IARPA is reconsidering the technical focus of Phase 3. Proposers are therefore encouraged to read the upcoming BAA carefully and to note that there may be differences between the BAA and the information contained in this briefing (see slide 3).
Integrated Cognitive-Neuroscience Architectures for Understanding Sensemaking

Proposers’ Day
January 19, 2010

Brad Minnery, Ph.D.
Program Manager, Office of Incisive Analysis
Intelligence Advanced Research Projects Activity
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 the Proposers’ Day changes requirements set forth in the 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/accepted for ICArUS

- Proposals will be due approximately 45 days after the BAA is published

- Take advantage of this time to start developing your ideas
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
**ICArUS Program (5 yrs)**

**Goal:** Computational cognitive neuroscience *models* that explain, predict and emulate human *sensemaking*

- **Explain** sensemaking based on underlying neuro-cognitive mechanisms
- **Predict** human sensemaking performance, including cognitive biases & failure modes
- **Emulate** human sensemaking on complex analysis tasks

**ICArUS Vision (5+ yrs)**

**Impact:** Incisive analysis *tools* for enhancing the performance of human-in-the-loop analysis systems

- An *analytic force multiplier*, taking over low-level sense-making tasks from overburdened analysts
- A ‘mirror brain’ to assist analysts in examining assumptions, identifying biases and re-examining underweighted or missed evidence
- A generator of "insight models" for facilitating analyst intuition & discovery
- A *predictive tool* for evaluating the potential impact of new analytic techniques, tools & methodologies

**Transition**

**Goal:** Computational cognitive neuroscience *models* that explain, predict and emulate human *sensemaking*

- **Explain** sensemaking based on underlying neuro-cognitive mechanisms
- **Predict** human sensemaking performance, including cognitive biases & failure modes
- **Emulate** human sensemaking on complex analysis tasks

**Impact:** Incisive analysis *tools* for enhancing the performance of human-in-the-loop analysis systems

- An *analytic force multiplier*, taking over low-level sense-making tasks from overburdened analysts
- A ‘mirror brain’ to assist analysts in examining assumptions, identifying biases and re-examining underweighted or missed evidence
- A generator of "insight models" for facilitating analyst intuition & discovery
- A *predictive tool* for evaluating the potential impact of new analytic techniques, tools & methodologies
**Definitions & Assumptions**

*Sensemaking* is the process of generating and evaluating hypotheses to explain data that is sparse, uncertain, and potentially deceptive.

“All individuals assimilate and evaluate information through the medium of…‘frames.’ These are experience-based constructs of assumptions and expectations both about the world in general and more specific problem domains.”

– from *Tradecraft Review. CIA Kent Center for Analytic Tradecraft*

Sensemaking entails:
- Fitting one or more explanatory *frames* (mental models) to the *data*
- Actively *seeking* additional data to confirm or refute the current frame/hypothesis
- *Evaluating* the quality of the data
- *Deciding* whether to accept/reject the current frame
- Continuously *learning* new frames & modifying existing frames

**ICArUS models must address all of the above functions**
Sensemaking in Intelligence Analysis

<table>
<thead>
<tr>
<th>Observe data</th>
<th>Learn frame</th>
<th>Apply frame</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data:</strong> Scenes, e.g. overhead images</td>
<td>Typical scene contents &amp; geometry</td>
<td>What type of facility is this?</td>
</tr>
<tr>
<td><strong>Spatial Context</strong></td>
<td></td>
<td>What is the function of these buildings?</td>
</tr>
<tr>
<td><strong>Relational network</strong></td>
<td></td>
<td>What is stored in this structure?</td>
</tr>
<tr>
<td><strong>Data:</strong> Actors &amp; interactions</td>
<td>Organization / command structure</td>
<td></td>
</tr>
<tr>
<td><strong>Frames</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Script</strong></td>
<td></td>
<td>What is his operational role?</td>
</tr>
<tr>
<td><strong>Data:</strong> Ordered events</td>
<td></td>
<td>Who does he report to?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What valuable info might he have?</td>
</tr>
</tbody>
</table>

- What type of facility is this?
- What is the function of these buildings?
- What is stored in this structure?
- What is his operational role?
- Who does he report to?
- What valuable info might he have?
Out of Scope

- Putting EEG caps on intelligence analysts
- Models of *group* cognition (ICArUS focus is on *individual* brains/minds*)
- *Isolated* models of single brain/cognitive systems (ICArUS focus is on *integrated* models involving *multiple* systems)
- *Non*-biologically inspired approaches (i.e. classical AI, pure machine learning)
- Neurophysiological and behavioral data collection (ICArUS is principally a *modeling* effort!)
- Visual object recognition
- Natural Language Processing
- New hardware development
- Tools & widgets (for visualization, collaboration, assisted reasoning, etc)
Recent advances in neuroscience research:
- new neural recording techniques
- new data analysis methodologies
- increased focus on computational modeling

Result: Proliferation of computational models describing how brain accomplishes:
- Learning and memory
- Attention
- Decision making
- Goal-directed behavior
- Sensory perception
- Multisensory integration
- And many etc’s…”

The scientific foundations are now sufficient to begin constructing an integrated neuro-cognitive model of human sensemaking.
ICArUS seeks to develop *computational cognitive neuroscience* (CCN) models of human sensemaking.

CCN is defined as an emerging discipline “at the intersection of neuroscience, cognitive psychology, and computational modeling, where neuroscience-based computational models are used to simulate and understand cognitive functions such as perception, attention, learning and memory, language, and [other] functions” -- [http://www.ccnconference.org/](http://www.ccnconference.org/)

- Required neural structures will be specified in the BAA (discussed later)
- *Integration* of individual brain systems into a unified architecture is essential
- Exact level of biological detail will be decided by the modelers; extraneous details (e.g. 3D neuronal microstructure) will be discouraged
- Models should focus on ‘higher-level’ cognition (attention, learning & memory, decision making) as opposed to ‘lower-level’ perceptual processes (e.g., visual feature extraction)
- Level of biological detail may vary among modeled brain areas due to inconsistencies in our understanding of brain function

**Teams are anticipated to be multidisciplinary, with collaboration among neuroscientists and theoreticians strongly encouraged**
## Required Neural Systems

<table>
<thead>
<tr>
<th>Brain systems</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefrontal Cortex</td>
<td>Attention, cognitive control, working memory, goal-oriented behavior, decision making</td>
</tr>
<tr>
<td>Parietal Cortex</td>
<td>Evidence integration, decision making, multimodal sensory representation, spatial reasoning, estimation of value and uncertainty</td>
</tr>
<tr>
<td>Medial Temporal Lobe, Hippocampus</td>
<td>Recognition and recall, declarative (episodic and semantic) memory, spatial cognition, relational processing, sequence learning</td>
</tr>
<tr>
<td>Basal Ganglia / Dopaminergic Systems</td>
<td>Reinforcement learning, reward signaling, slow statistical learning, action sequencing, procedural learning, decision making</td>
</tr>
<tr>
<td>Anterior Cingulate Cortex</td>
<td>Error signaling, cognitive control, conflict monitoring, decision making</td>
</tr>
<tr>
<td>Brainstem Neuromodulatory Systems</td>
<td>Attentional arousal, transition between exploitative and exploratory behavioral modes</td>
</tr>
<tr>
<td>Amygdala, Orbitofrontal Ctx, Limbic Structures</td>
<td>Emotional arousal, decision making, estimation of value</td>
</tr>
</tbody>
</table>
**Program Structure**

**Modeling Teams**

**Responsibilities:** Construct models; work with T&E team to perform model validation / testing.

- Team 1
- Team 2
- Team n

**Program Management**

**Test & Evaluation Team (TBD)**

**Responsibilities:** Develop T&E framework (Challenge Problems and Cognitive Fidelity Assessments). Validate model performance vs human performance

**Panel of Neuroscience Experts (Govt / FFRDC)**

**Responsibilities:** Conduct neural fidelity assessments
Challenge Problems

**Purpose:** Provide an integrated task environment / test framework for comparing model performance to human performance on end-to-end sensemaking tasks

- Challenge Problems will:
  - incorporate, in an integrated fashion, *all* major sensemaking processes within a single task framework
  - be developed by the independent Test & Evaluation Team in consultation with Modeling Teams and intelligence analysts
  - balance the need for operational relevance/realism and scientific rigor
- The following slides provide a *notional* description of the CP framework meant to help guide proposal development
- Details are subject to change
- Complete test specification & data sets will be developed/released *during* the course of the program
Challenge Problems will involve the analysis of simulated Geospatial Intelligence (GEOINT) data

GEOINT: “the exploitation and analysis of...geospatial information to describe, assess, and...depict physical features and geographically referenced activities on the earth. GEOINT consists of...imagery intelligence and geospatial information” (Title 10 U.S. Code 467).

Models will **not** be required to:
- Process raw imagery (pixels)
- Perform visual feature extraction / object recognition
- Process natural language
Additional Challenge Problem Characteristics

- Escalate in complexity as program progresses from Phase 1 → Phase 3
- Different for each Phase but similar in overall format (GEOINT)
- Involve “directed” sensemaking tasks (model must answer specific questions)
- Difficulty level is challenging for humans
- Inputs: multiple GEOINT ‘data layers’
  - each layer a separate info modality
  - multiple time points per layer (discrete time)
- Input format: multi-dimensional feature vector
- Questions are multiple choice
- Output format: confidence estimates for the different answer options, response selection
Challenge Problems: Inputs

**What an analyst sees**
(layers of GEOINT data)

**ICArUS inputs**
(Multi-d feature vector / nXn grid)

- **IMINT**
- **GMTI**
- **SIGINT**
- **SIGACTs**
Challenge Problems: Measuring Performance

Models will be compared with humans performing same Challenge Problem tasks

- Inherently rich and extensible task framework
- Challenge Problems will be designed to control for humans’ advantage in background knowledge

*Bars show changing confidence levels for response alternatives A-D as more data layers are revealed*
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
Program Phases

- Distinguished principally by:
  - minimum required *capabilities of model*
    - types of frames learned/applied
    - search & decision processes
  - structure and complexity of *task environment*
    - role of *time*
    - probabilistic structure
  - *metrics & milestones* (to be discussed later)

Teams will be encouraged/rewarded for aggressive approaches that exceed the minimum targets.
Phase 1 (24 months)

**Objective:** Construct *integrated* neuro-computational model that captures *all* core sensemaking functions and that successfully performs the Phase 1 Challenge Problem

**Minimum capabilities of model:**
- Process *spatial* input data
- Operate in probabilistically *constant* environment
- Learn & apply *spatial context frames* to perform basic inferences
- Demonstrate *simple decision making* (e.g., select relevant data layer/source)

---

**Data Layers**

- A
- B
- C
- D
- E
- F
- G
- H
- I

**Frames**

**Spatial contexts:** object/event membership and spatial relationships within a scene

**Ex:** Spatial layout of facilities (utility substations, factory-warehouses, etc); spatial patterns of human activity (traffic densities, congregation sites, IED events)

---

**Probabilities (Constant)**

- - -
Phase 2 (18 months)

**Objective:** Expand functionality of model to include (at minimum) the ability to:

- Process *temporal* input data
- Operate in and adapt to probabilistically *changing* environment
- Learn & apply *event sequence frames* in conjunction with spatial context
- Demonstrate *complex* decision making (e.g., select relevant data layer / time slice)

---

**Data Layers**

**Frames**

---

**Event sequences:** a.k.a. *scripts*: recurring actions and events ordered in time

**Ex:** vehicular traffic patterns; industrial activity (e.g. temporal patterns in factory emissions spectra, construction sequences), crowd movements, comms patterns
**Objective:** Extend the ability of the models to perform sensemaking under situations in which data (or the absence thereof) may be the result of *denial and deception*.

**D&D tactics:** e.g. misdirection, concealment. Plus: awareness of own frames/sensemaking strategies

**Ex:** Change in location of facilities, weapons caches, etc; introduction of deliberately misleading evidence
## Three Phases

<table>
<thead>
<tr>
<th>Phase</th>
<th>frames(^1)</th>
<th>time(^2)</th>
<th>statistics of environment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase 1</strong></td>
<td><strong>spatial context</strong></td>
<td>no</td>
<td>constant</td>
</tr>
<tr>
<td><strong>Phase 2</strong></td>
<td><strong>spatial context, event sequences</strong></td>
<td>yes</td>
<td>changing</td>
</tr>
<tr>
<td><strong>Phase 3</strong></td>
<td><strong>spatial context, event sequences, denial &amp; deception</strong></td>
<td>yes</td>
<td>changing-<em>adaptive</em></td>
</tr>
</tbody>
</table>
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
Three classes of metrics aimed at answering two different questions

1) Neural Fidelity and 2) Cognitive Fidelity

**How?**
Do the models solve problems using the same processes and mechanisms that humans do?

- **neuro-fidelity (brain models)**: yes/no
- **cognitive-fidelity (cognitive biases)**: yes/no

3) Comparative Performance

**How well?**
How do models compare with human performance on sensemaking Challenge Problems?

- model performance (relative to humans): % performance
Purpose: Ensure that teams remain faithful to the ICArUS Program’s objective of building brain-based cognitive models

Assessments will be made by PM Team with guidance from independent Panel of Experts (Govt, FFRDC). Judgments will be based on analysis of technical reports, model source code, and models’ activation dynamics during task execution.

Criteria: For each of seven key brain systems…

- Does model incorporate neurobiologically plausible components & design principles?
- Does the model maximally exploit existing knowledge of brain’s functional architecture?
- Are model’s internal dynamics during task execution consistent with the literature?

Target Metrics: (fraction of key brain areas faithfully represented):
Phase 1: 3 of 7
Phase 2: 5 of 7
Phase 3: 7 of 7
**Purpose:** Assess whether ICArUS models faithfully capture key biases and other cognitive idiosyncrasies known to impact (often detrimentally) human sensemaking

- Cognitive Fidelity Assessments will be:
  - conducted within same test environment as overall Challenge Problems
  - performed using *complete* integrated model
  - performed using consistent parameter settings *(no tweaking the knobs for each individual test!)*
Examples of cognitive biases / idiosyncrasies of interest to ICArUS

- Confirmation bias
- Anchoring and adjustment
- Inattentional blindness
- Change blindness
- Satisfaction of search
- Representativeness
- Availability
- Vividness
- Probability matching
- Inductive biases (*learning to learn*)
- Overconfidence effect
- Over-reliance on evidence labeled ‘high value’ (e.g. “classified”)

**Target Metrics:** (fraction of key cognitive biases exhibited by model):
Phase 1: 2 of 4
Phase 2: 5 of 8
Phase 3: 8 of 12
Metrics: Cognitive Fidelity

- Based on a 3-way comparison of model/human/normative responses.
- Key question: Does model deviate from normative behavior in same way as human?
- Separate assessments will be conducted for each bias of interest, with Pass/Fail result for each bias.

Data layers (inputs)

Responses (outputs)

Model deviates from normative response via same pattern as human.
Purpose: Assess models’ ability to emulate humans’ overall performance on challenging sensemaking tasks.

Approach: Compare item-by-item response patterns of models to those of humans performing same Challenge Problem tasks. Measure the average divergence between model vs human responses.

Goal is to match humans’ response patterns – not just their scores.
Metrics: Comparative Performance

Questions:

- i = 1
- i = 2
- i = 3
- etc.

Data layers:

Human Model Human Model Human Model Human Model

Human Model Human Model Human Model Human Model

Entropy

- KLD

Base Entropy $E$, Cross Entropy $E'$, $\text{KLD} = E' - E$
Target Metrics for *Comparative Performance* (Model v Human)

<table>
<thead>
<tr>
<th>Phase</th>
<th>% Match (Model-Human)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>50%</td>
</tr>
<tr>
<td>Phase 2</td>
<td>65%</td>
</tr>
<tr>
<td>Phase 3</td>
<td>80%</td>
</tr>
</tbody>
</table>

% of model theoretical max
## Metrics: Summary

<table>
<thead>
<tr>
<th>Test</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neural Fidelity</strong></td>
<td>3/7</td>
<td>5/7</td>
<td>7/7</td>
</tr>
<tr>
<td>(fraction of key brain areas faithfully represented)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cognitive Fidelity</strong></td>
<td>2/4</td>
<td>5/8</td>
<td>8/12</td>
</tr>
<tr>
<td>(fraction of cognitive biases exhibited)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparative Performance</strong></td>
<td>50%</td>
<td>65%</td>
<td>80%</td>
</tr>
<tr>
<td>(% of human performance level)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Timeline

<table>
<thead>
<tr>
<th>Task</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6 mo</td>
<td>12 mo</td>
<td>18 mo</td>
</tr>
<tr>
<td></td>
<td>24 mo</td>
<td>30 mo</td>
<td>36 mo</td>
</tr>
<tr>
<td></td>
<td>42 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Select/develop component models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrate models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 1 Test &amp; Eval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 2 Test &amp; Eval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 3 Test &amp; Eval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Challenge Problem(s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collect Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administer Phase 1 tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administer Phase 2 tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administer Phase 3 tests</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Technical Exchange Mtgs to be held at 6-month intervals. Technical/Financial status reports due monthly.
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
Award Plan

- **5-yr program**
  - Phase 1 – 24 months (12-month Base + 12 month Option)
  - Phase 2 – 18 months (Option)
  - Phase 3 – 18 months (TBD)

- Criteria for advancing to next phase: sufficient progress in current phase metrics
- Number of awards depends upon:
  - Quality of the proposals received
  - Availability of funds
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
Eligibility

- All proposals must address all facets of the program
- Teaming/collaborations are strongly encouraged
  - Networking and team formation is responsibility of the proposers
- Foreign organizations and/or individuals may participate
  - Must comply with Non-Disclosure Agreements, Security Regulations, Export Control Laws, etc, as appropriate
Eligibility

The following are NOT eligible to submit proposals to this BAA or participate as team members under proposals submitted by eligible entities.

- 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
Other Issues

  
  Example: “any instance where an offeror, or any of its proposed subcontractor teammates, is providing either scientific, engineering and technical assistance (SETA) or technical consultation to IARPA.”

- Publication is encouraged, but...

- Performers should provide a pre-publication soft copy to:
  - IARPA ICArUS Program Manager
  - Contracting Officer’s Technical Representative
Outline

- Program Overview
- Program Phases
- Program Metrics & Milestones
- Award Information
- Eligibility Information
- Application Review Information
Application Review Information

In descending order of importance

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

Thanks to:
- Dr. Anthony Boemio
- Mr. Kevin Burns
- Dr. Peter Highnam
- Rest of IARPA Front Office
- Events Team
Dr. Brad Minnery
Program Manager
IARPA Office of Incisive Analysis
Office of the Director of National Intelligence
Intelligence Advanced Research Projects Activity
Washington, DC 20511

Phone: (301) 851-7449
Email: dni-iarpa-baa-10-04@ugov.gov
(include IARPA-BAA-10-04 in the subject line)

Website: www.iarpa.gov
Questions?
Thank You!