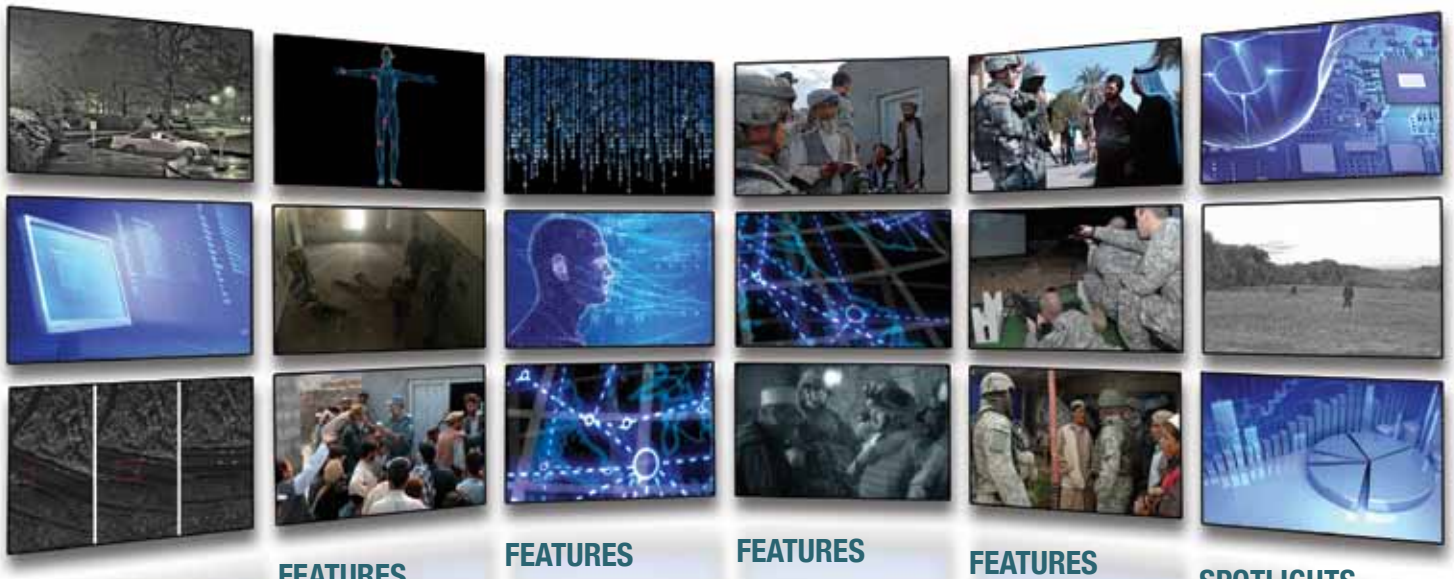
DEPUTY UNDER SECRETARY OF DEFENSE FOR SCIENCE AND TECHNOLOGY

HSCB

HUMAN SOCIAL CULTURE BEHAVIOR MODELING PROGRAM





# FEATURE ARTICLE SOCIAL RADAR

**By Mark Maybury, PhD**

## Social Sensing

**In the 20th century, advances in radar, sonar, and infrared sensing dramatically improved our ability to perceive physical objects through air, water, and darkness/camouflage. In the 21st century, the strategic center of gravity has extended beyond these traditional kinetic targets, and US national security leaders are increasingly emphasizing "hearts and minds" as critical determinants of our strategic success or failure. Accordingly, today's diplomats, developers, and defenders need a set of capabilities that will provide insight into the attitudes, perceptions, and intentions of international citizens and leaders. They need a social radar that will provide situational awareness and decision support for strategic communication, countering violent extremism, and building partnership capacity.**

## Requirements for a Social Radar System

Conventional access to foreign public beliefs and opinions most frequently occurs via polling or focus groups. While extremely valuable, these are manually intensive, expensive, episodic, and subject to interviewer bias and interpretation error. Although not without challenges, automated, large scale, continuous analysis of global communications promises to provide relatively inexpensive, wide area, multidimensional, persistent social sensing.

For social radar to be as revolutionary as conventional physical sensors, it must exhibit some fundamental system properties. First, it must provide a *global perspective* that includes persistent, worldwide, geo-located, real-time capture and analysis of indicators to include areas with limited connectivity, denied access, or active censorship. This implies that social radar will be *multilingual* and *multicultural*, supporting transcription, summarization, translation, and interpretation across languages and societies. Moreover, because of the multifaceted nature of communication, social radar must be *multimodal*, providing the ability to process multiple media (e.g., radio, television, newspapers, websites, blogs, wikis) and multiple modalities (e.g., text, audio, imagery, action). This presents a number of technical challenges, including text understanding, speech recognition, and image and video understanding. By its nature, social radar should detect and track *social* interactions among individuals, groups, tribes, and societies, using direct and indirect indicators to sense perceptions, attitudes, beliefs, opinions, and behaviors, as well as social network roles, and relationships. It will need to provide

CONTINUED ON PAGE 5

# Welcome

*Welcome to the sixth edition of the HSCB newsletter. This issue has a special focus on the importance of HSCB data. Four experts in this area have each written a feature article discussing data. The first article addresses the need for social radar to provide situation awareness, support strategic communication, and build partnership capacity. The second article focuses on the Bayesian inference—a popular approach to evidence-driven hypothesis analysis which measures how an initial probability in the hypothesis may change when evidence relating to it is observed. Next, we focus on political analysis using automated event data and sentiment coding. The final article discusses the challenges associated with understanding survey data and its related complexities.*

*This edition also continues to follow many topics we introduced in the last issue of the HSCB newsletter. Expanding on the interview I gave in the spring newsletter edition regarding Small Business Innovation Research (SBIR) topics, this issue discusses the SBIR review process and dives deeper into the topic areas and Phase I awards coordinated by CDR Joseph Cohn.*

*We also continue to look at the HSCB Program's DoD-wide "brother and sister programs" with a special emphasis on the Air Force Research Laboratory (AFRL). Starting in FY 2008, three AFRL Directorates joined forces to embark on a common goal, Predicting Adversary Behaviors. In support of this new initiative, AFRL divided the problem into two areas: Understanding the Operational Environment and Understanding the Adversary. This newsletter describes leading AFRL tools currently in development that address each of these areas. And finally, we look deeper at the last two hard research challenges previously presented in the spring newsletter with spotlights on the work of Dr. Keith Gremban and Dr. Allison Abbe.*

*In the next edition of this newsletter, we plan to offer feature articles on mission analysis and planning as follow-up to the important issue of data. We will also highlight the 2010 AHFE 3rd International Conference on Applied Human Factors and Ergonomics which runs jointly with the 1st International Conference on Cross-Cultural Decision Making in Miami, Florida. I look forward to seeing many of you there.*

***CAPT Dylan Schmorrow, MSC, USN, PhD***
*Acting Director, Human Performance, Training and BioSystems Research Directorate*
*Office of the Director, Defense Research & Engineering*
*Office of the Secretary of Defense*

## FEATURE ARTICLE

# A NON-BAYESIAN APPROACH TO EVIDENCE-DRIVEN HYPOTHESIS ANALYSIS AND THE TOPOLOGICAL HYPOTHESIS ANALYSIS TOOL (TOPHAT®)

**By Paul Garvey, PhD**

## Introduction

A popular approach to evidence-driven hypothesis analysis is Bayesian inference. Bayesian inference originates in statistical theory. It is a way to measure how an initial probability in the truth or falsity of a hypothesis may change when evidence relating to it is observed. For some, the use of probability in this context is undesirable. Its accuracy as a measure of belief in the trueness of a hypothesis is impossible to prove. The main difficulty often comes in assessing inputs primary to a Bayesian inference model—dual judgmental probabilities such as "What is the chance this evidence would be seen if the hypothesis were true?" and "What is the chance this evidence would be seen if the hypothesis were not true?" The accuracy of probability as a measure of belief in the trueness of a hypothesis is impossible to prove. Even if its accuracy could be ascertained, the nature of probability is such that an event with a high chance of occurrence might not occur.

In practice, decision makers look for measures indicative of the strength in the veracity of a hypothesis. Probability is not the only such measure. Value function theory offers an option to probability theory for evidence-driven hypothesis analysis. Value function theory is a formalism that enables judged preferences in the performance of options across evaluation criteria to be expressed by mathematical functions that capture the goodness of each option, over the levels of each criterion. These functions are known as *value functions*. They are measurable value functions when they capture one's preference ordering and strength of preferences between each criterion's levels. When value function theory is applied to evidence-based hypothesis analysis, *strength of trueness* is defined as a measure analogous to strength of preference. Strength of trueness represents

**STAFF**

**Published by:**
Strategic Analysis, Inc.
4075 Wilson Blvd.
Suite 200
Arlington, VA 22203

**Layout and Design:**
Amy Cauffman

**Technical Editor:**
Becky Bortnick

# A NON-BAYESIAN APPROACH TO EVIDENCE-DRIVEN HYPOTHESIS ANALYSIS AND THE TOPOLOGICAL HYPOTHESIS ANALYSIS TOOL (TOPHAT®)

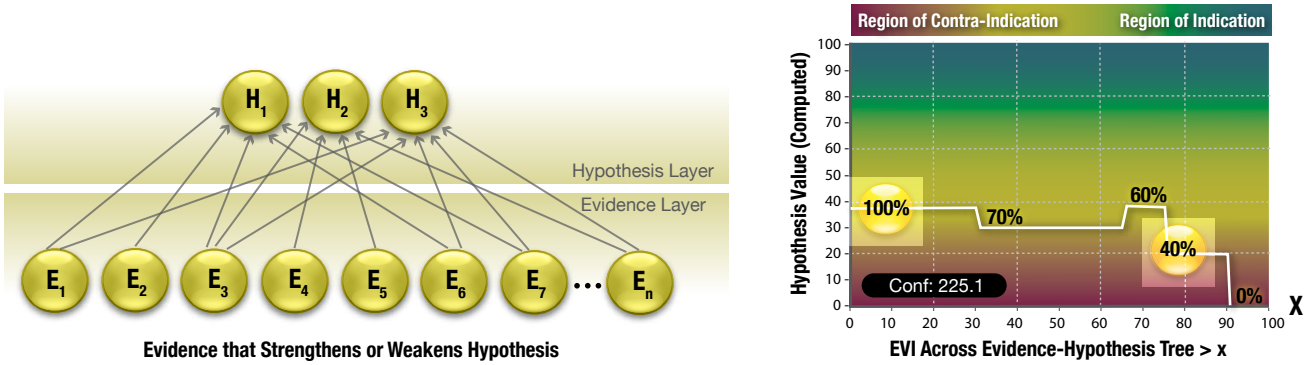



**Figure 1. Hypothesis Response to Changing Evidence Varacity**

judgment in the inferential *force of evidence* and the *direction of evidence* for or against the hypothesis or alternative hypotheses (Shum, 1994).

topHAT® is a newly developed software application that applies value function theory and Euclidean decision algorithms to analyze the veracity of competing or non-competing hypotheses. It does this without the need to directly probabilistically reason about evidence. The analytic protocols in topHAT® are non-Bayesian and built on the axioms of (von Neumann and Morgenstern, 1944) and (Keeney and Raiffa, 1976) decision theory. topHAT® provides decision makers a full trace and measure of the diagnosticity of all pieces of contributing evidence. This can be done according to the source basis, integrity, and veracity of evidence. Measuring evidence diagnosticity has been an elusive metric in the analytic community. In topHAT®, algorithms used to measure evidence diagnosticity derive from the entropy metric from Shannon information theory (Shannon and Weaver, 1998).

## Summary

The use of value function theory in evidence science offers important contributions to *sense-making* in understanding the complexities of human, social, cultural, and behavioral interactions. At best, the methods herein inform policy considerations and offer ways to tradeoff the efficacies of policy options. At worst, consumers of products from these methods view them with unwarranted predictive certainty.

Hypothesis analysis is an interpretive and judgment-driven exercise. As such, decision makers should not expect pinpoint predictions. Instead, they should expect (1) insight into the dynamics of a social-political-economic problem that underlies the analysis of a hypothesis, (2) an analysis of the effects of complex interactions between numerous qualitative and quantitative variables on the indicative trueness of a hypothesis, and (3) understanding the sensitivity of outputs to changing inputs.

A final consideration is the importance of looking for evidence that disconfirms rather than confirms a hypothesis. Confirmation bias is a well-known issue in behavioral decision theory. It refers to an innate preference for selectively seeking information that corroborates potentially prejudiced notions about a hypothesis. For this reason, the value function should be two-sided with one region for confirming evidence and one region for disconfirming evidence —with respect to grading the force of evidence on a hypothesis. We should never truly rest in thinking we have confirmed with certainty a hypothesis, but continually challenge the merits and basis for the strength with which it is indicated. We close with an eloquent commentary on this topic from Richards Heuer:

*"Apart from the psychological pitfalls involved in seeking confirmatory evidence, an important logical point also needs to be considered. The logical reasoning underlying the scientific method of rejecting hypotheses is that "...no confirming instance of a law is a verifying instance, but that any disconfirming instance is a falsifying instance." In other words, a hypothesis can never be proved by the enumeration of even a large body of evidence consistent with that hypothesis, because the same body of evidence may also be consistent with other hypotheses. A hypothesis may be disproved, however, by citing a single item of evidence that is incompatible with it." (Heuer, 1999)*

To learn more about the non-Bayesian approach to evidence-driven hypothesis analysis described herein, as well as the analytics and algorithms created for topHAT®, contact the author at pgarvey@mitre.org. ♦

## References

Garvey, P.R. "A Non-Bayesian Approach to Evidence-Driven Hypothesis Analysis", MITRE Paper, The MITRE Corporation, June 2010.

Heuer, Richards J. *Psychology of Intelligence Analysis.* Langley: Center for the Study of Intelligence, The Central Intelligence Agency, 1999.

Keeney, Ralph L., and Howard Raiffa. *Decisions with Multiple Objectives Preferences and Value Tradeoffs.* New York: John Wiley & Sons Inc., 1976.

Shannon, Claude E. and Warren Weaver. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press, 1998.

Shum, David A. *The Evidential Foundations of Probabilistic Reasoning.* Evanston: Northwestern University Press, 1994.

von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

# FEATURE ARTICLE INSTANT DATA, JUST ADD DICTIONARIES: POLITICAL ANALYSIS USING AUTOMATED EVENT DATA AND SENTIMENT CODING

**By Philip Schrodt, PhD and Steve Shellman, PhD**

Event data are "day-by-day coded accounts of who did what to whom as reported in the open press," (Goldstein, 1992, 369). Most basic event datasets assign categorical codes for the actor taking the action, the target receiving the action itself (for example a bombing or an agreement), and the date of the action. Other information such as location, political context, and number of individuals (casualties, demonstrators) associated with the event might also be coded.

Event data have three distinct advantages over other forms of social science data. First, they focus on actors who are making specific, time-dependent and inter-dependent tactical and strategic decisions (See Shellman, 2008; Shellman et. al. 2010). In contrast, most other international social science data focus on structural attributes of countries or dyads (e.g. population, GDP, whether a dyad is at a war, level of trade). While structural variables can affect behavior, event data directly *measure* behavior and can be used to infer attitudes as well as allowing us to study the strategic inter-dependence of actors' choices.

Second, event data track behavioral processes over time at a high level of granularity. Most international political data are aggregated by year, but researchers and analysts are often more interested in processes that change much more quickly. An insurgency, peace process, or international negotiation will often change from week-to-week or month-to-month. Event data can detect rapid changes, and in a wide variety of academic and policy-oriented analyses have been shown to accurately forecast such changes in actors' behavior.

Finally, event data are now very inexpensive to update in near real time once the relevant infrastructure is in place. Until recently, event data were coded manually—a very expensive, error-prone and time-consuming process—but now a number of specialized coding systems, including Penn State's open source TABARI, SAE's Xenophon, Virtual Research Associates VRA Coder, and Social Science Automation's BEN can produce event data using automated methods. Moreover, studies show that machine coded events data are as accurate as human coded data (King and Lowe, 2003). With automated coding, the coding rules are transparent, the data are easily and quickly reproducible, the data can be regenerated using alternative coding schemes, and the data are unaffected by individual coders' biases. Moreover, automated coding dramatically reduces the time required for coding once the input texts have been formatted and coding dictionaries prepared. For example, we can recode the 8-million stories covering 12 years of stories for 29 countries currently used for DARPA's ICEWS project in a few hours, and updates to that data can be made in minutes.

The automated systems, however, do require dictionaries of noun and verb phrases in order to identify actors and events. These dictionaries were generated manually in the past, but we are currently developing automated techniques to populate actor dictionaries from the texts being coded, and to identify possible missing phrases in the verb dictionaries. The manual downloading of stories, meanwhile, is likely to be replaced by the automated acquisition of texts from web sites and RSS feeds, at least for near-real-time coding. These developments will further increase the overall efficiency of the data generation process.

**THE RISE OF BLOGS AND INTERNET SITES SUCH AS TWITTER, FACEBOOK, AND MYSPACE HAS FUELED THE ELECTRONIC EXPRESSION OF EMOTION, WHICH MIGHT ALSO BE TURNED INTO HARD DATA.**

In addition to coding events, we are also beginning to code sentiment data. The rise of blogs and internet sites such as Twitter, Facebook, and MySpace has fueled the electronic expression of emotion, which might also be turned into hard data. Many theories of conflict, insurgency, and counter-insurgency stress the necessity of winning the hearts and minds of the masses. While feelings are more difficult to code than occurrences, a number of techniques are now available that can capture this information and use it to project future behavior, much as it is possible to predict the likely success of a movie based on the emotions expressed in blogs (Sadikoy, Parameswaran and Venetis, 2009). We surmise that automated sentiment data will play a large role in defense and social science related projects such as the DARPA ICEWS project and Combatting Terrorism Technical Support Office (CTTSO) HSCB Scalable Modeling System Prototype now and in the future, replacing survey data with automated measures that can be produced faster, better, and cheaper. ♦

## References

Goldstein, Joshua S. "A Conflict-Cooperation Scale for WEIS Events Data." Journal of Conflict Resolution 36 (1992): 369-385.

King, Gary and Will Lowe. "An Automated Information Extraction Tool For International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design" International Organization Vol. 57, No. 03 (2003): 617-642.

Sadikov, Eldar, Aditya Parameswaran, and Petros Venetis. "Blogs as Predictors of Movie Success." 2009. In the 3rd International AAAI Conference on Weblogs and Social Media.

Shellman, Stephen M., Clare J. Hatfield, and Maggie J. Mills. "Disaggregating Actors in Intranational Conflict." Journal of Peace Research 47 (2010): 83-90.

Shellman, Stephen M. "Coding Disaggregated Intrastate Conflict: Machine Processing the Behavior of Substate Actors Over Time and Space." Political Analysis Vol. 16, Issue 4 (2008): 464-477.

# FEATURE ARTICLE STATISTICAL ANALYSIS, SURVEYS, AND CATEGORICAL DATA[1]

**By Jeffrey Appleget, PhD**

Since Irregular Warfare (IW) is all about influencing the local population, the difficulty for decision makers in any IW scenario is finding methods for measuring the impact of operations on the local population. One instrument that has been used both in Iraq and Afghanistan to understand the population is surveys. The challenge with understanding survey data is that many survey questions yield categorical data. While there are statistical analysis techniques for categorical data, these techniques are typically more complex than the techniques used to analyze continuous (numerical) data.

We are familiar with the mean, a statistic that measures the central tendency for numerical data, and probably remember that the standard deviation characterizes the spread of numerical data. So by using two statistics, mean and standard deviation, we get a fair first summary of a numeric variable and gain an understanding of its features.

But what about categorical data? Categorical data can be distinguished by type: nominal (no natural ordering, e.g., eye color), ordinal (has ordered levels e.g., a Likert scale: strongly agree, agree, neither agree or disagree, disagree, strongly disagree) and interval (data that has numeric distances between any two levels e.g., annual income ranges such as tax tables). With ordinal data and interval data, often a numeric score is assigned to each level (e.g. 1=strongly agree, 2= agree, etc.) and then analysis is conducted as if each response is quantitative. The difficulty with this is that for these types of categorical data, there is no one natural set of scores to assign. We might assign scores 0, 10, 11, 12, 20 to the Likert levels, rather than equally spaced scores 1, 2, 3, 4, 5. By careful thought of which scores to assign we can often gain some understanding of ordinal (and interval) data by applying analysis techniques for quantitative data. At other times, when it is not possible to assign a reasonable set of scores, it is more useful to ignore the natural ordering of ordinal and interval data and treat it as nominal.

If we want to use several predictors from a particular survey (explanatory or independent variables) to see if they can be grouped together to explain a particular response (response or dependent variable), we are now talking about fitting a model to the data. The simplest model most of us have seen fitted is a linear regression model, often known as "least-squares," where all the variables (explanatory and response) are numeric. Linear regression is easily extended to contend with both numeric and categorical explanatory variables. However, when one wants to fit a model where the response variable is categorical, there are various Generalized Linear Models (GLMs) that can be used. The particular GLM to be used is determined by the type of response variables. Logistic regression, loglinear models, and multinomial response models are all examples of GLMs used that can contend with categorical response variables. Logistic regression is used if the response is a categorical variable with only two levels (e.g. does the population support the current government or not). Multinomial response models are used if the categorical response takes more than two levels (e.g., do members of a population: support the government, are neutral, or support an insurgent group). As with any useful modeling technique, understanding the type of data one has, choosing the appropriate explanatory and response variables, and fitting the right GLM will require a little art thrown in with a lot of statistical science. ♦

1. Many of the concepts in this article come from "Categorical Data Analysis," by Alan Agresti, John Wiley and Sons, 1990.

CONTINUED FROM PAGE 1

# FEATURE ARTICLE SOCIAL RADAR

insight into *multiple domains* including politics (e.g., governance), economics, military/law enforcement (including crime and corruption), society, healthcare, education, and the environment. It should also be *passive* and *anonymous* to ensure signal validity and conceal queries which could otherwise bias respondents. Finally, it should preserve individual *security* and *privacy*.

## Technical Challenges: Sources and Methods

To detect, model, and forecast a broad range of phenomena, social radar will rely upon a rich set of sources including but certainly not limited to broadcast media, social media, social networking sites, and specialized data sources in areas such as health, economics, governance, or security. As with all sensors, social radar will need to be calibrated and have signatures developed to detect and track a broad range of phenomena. Because signals will arrive from myriad sources of various pedigrees, correlation and integration of evidence from social "signatures" as well as advanced modeling will be necessary to enable effective analysis and forecasting under uncertain information. Algorithms and methods will be needed to remove irrelevant or duplicative signals that clutter, interfere or obfuscate even a clean signal and thus impede sensemaking. Modeling of the various "terrains" (e.g., economic, technical, political, environmental, social) promises to assist in developing countermeasures to clutter so that extraneous "returns" (i.e., passive interference) can be eliminated.

Realizing such a grand vision of social sensing and tracking and the creation of a social radar will require the engagement of a broad set of stakeholders from a range of institutions including government, academic, industrial, non-governmental organizations (NGOs), and Federally Funded Research and Development Centers (FFRDCs). While daunting, social radar could be nothing less than revolutionary. ♦

**By John Salerno, PhD**

## Description of NOEM and SITA

Starting in FY2008, Air Force Research Laboratory (AFRL) Directorates (Information Directorate (RI), Human Effectiveness Directorate (RH) and Air Force Office of Scientific Research (AFOSR)) joined forces to embark upon a common goal, Predicting Adversary Behaviors. In support of this new initiative, AFRL looked towards publications by Joint Intelligence Preparation of the Operational Environment (JIPOE) and similarly divided the problem into two areas: (1) Understanding the Operational Environment, and (2) Understanding the Adversary. In this article we discuss leading AFRL tools currently in development that address each of these two areas.

The National Operational Environment Model (NOEM) is a strategic analysis/assessment tool that provides insight into the complex state space (as a system) that is today's modern operational environment. NOEM supports baseline forecasts by generating plausible futures based on the current state. It supports what-if analysis by forecasting ramifications of potential "blue" actions on the environment. NOEM also supports sensitivity analysis by identifying possible pressure (leverage) points in support of the Commander that resolve forecasted instabilities, and by ranking sensitivities in a list for each leverage point and response. NOEM can be used to assist decision makers, analysts and researchers in understanding the inner-workings of a region or nation state, the consequences of implementing specific policies, and the ability to implement new operational environment theories and models as they mature.

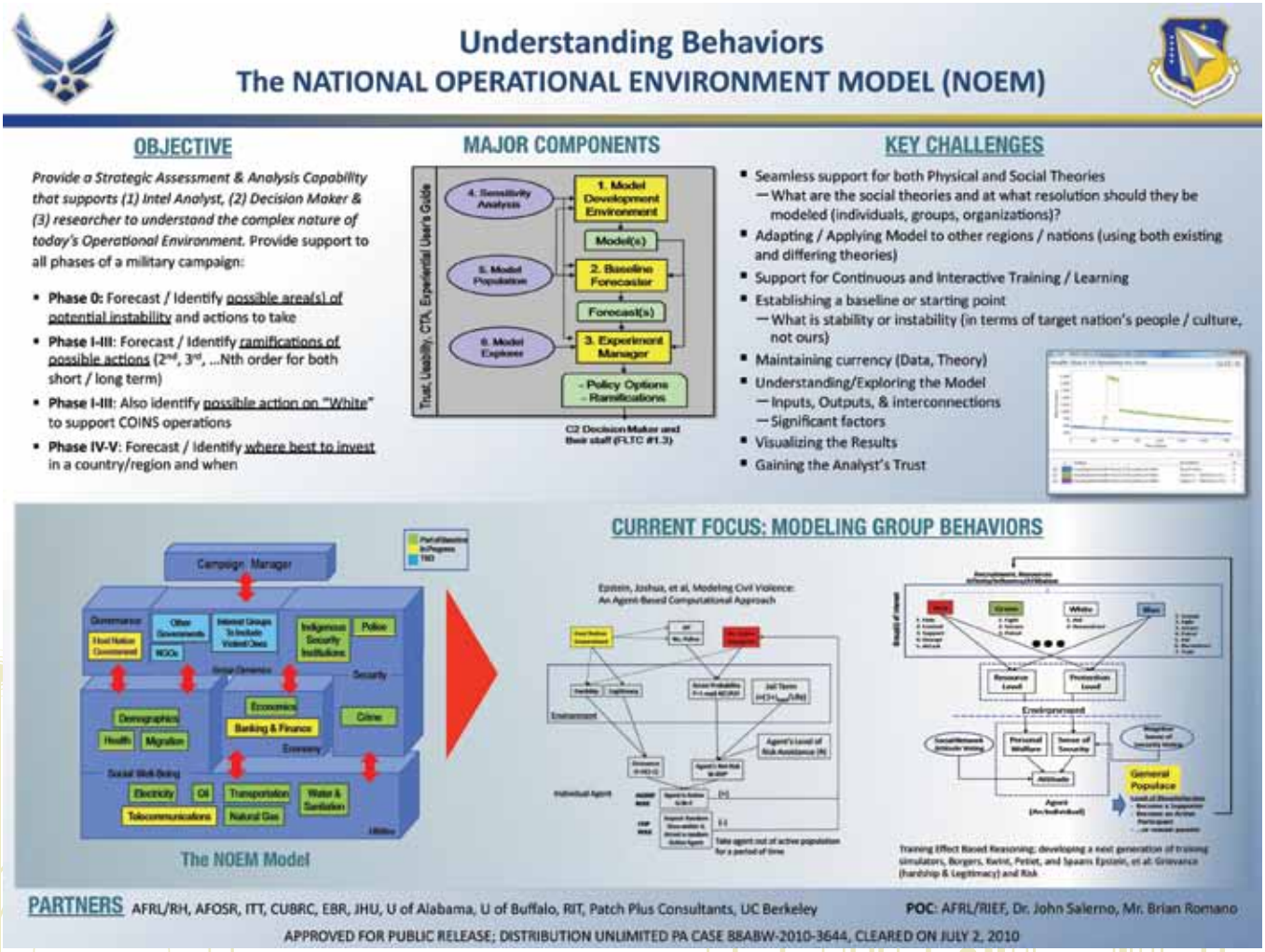



**Figure 1. Understanding Behaviors—The National Operational Environment Model (NOEM)**

NOEM is built upon an open-source license-free set of capabilities, and aims to provide support for pluggable modules that make up a given model. The architecture of NOEM consists of three major components: (1) Model Development Environment, (2) Baseline Forecaster, and (3) Experiment Manager. The heart of the NOEM is the model. It is composed of several modules (as defined using stability operations theory) that depict the various pillars of a nation-state, which are carefully integrated together to ensure that input/output dependencies are maintained between the modules. The Model Development Environment allows for the creation of new models or integration/application of existing models to new regions or nations. The Baseline Forecaster (along with the Model Population Subsystem) aims to maintain data currency by populating a data repository and to provide current and past snapshots/forecasts in time for the region/nation of concern. The Experiment Manager provides an analytical capability to exercise the model, allow what-if analysis, and a plug-in environment that allows for easy integration of future advanced analysis tools.

**THE FOCUS THIS YEAR IS ON MODELING THE SOCIAL AND BEHAVIORAL ASPECTS OF AN ENVIRONMENT, PRIMARILY THE FORMATION OF VARIOUS INTEREST GROUPS, THEIR BELIEFS, THEIR REQUIREMENTS, THEIR GRIEVANCES, THEIR AFFINITIES, AND THE LIKELIHOOD OF A WIDE RANGE OF ACTIONS, DEPENDING ON THEIR PERCEIVED LEVEL OF SECURITY AND HAPPINESS.**

NOEM currently has an extensive number of completed modules including: economic, security and social well-being pieces (e.g. critical infrastructure) along with a number of tools to exercise them. It provides the user with both a Turbo Tax® like interface for data entry/modification as well as an infrastructure toolkit to allow for geo-positioning of critical infrastructure assets that automatically encapsulate the code necessary to generate an effective model. The focus this year is on modeling the social and behavioral aspects of an environment, primarily the formation of various interest groups, their beliefs, their requirements, their grievances, their affinities, and the likelihood of a wide range of actions, depending on their perceived level of security and happiness. As such, several research efforts are currently underway to model human behavior from a group perspective, in pursuit of eventual integration and balance of populace needs and demands with their respective operational environments and their capacity to meet those demands.

The second effort, Situation Identification & Threat Assessment (SITA) provides the community with a capability that identifies potential threats and impacts that an adversary can have/create blue assets. The system takes a series of observed events, identifies a set of possible futures based on a *priori* of knowledge, applies constraints of what we know about ourselves (Knowledge of Us), and what we know about the adversary (Knowledge of Them), to then narrow down the possible futures to plausible futures. Each plausible future is ranked by the anticipated impacts/threats and a list of information requirements is identified that would help differentiate between all possible futures as well as the one that is actually unfolding.

**A KEY CHALLENGE WILL BE DEFINING AND UNDERSTANDING WHAT ARE THE INTENT/GOALS OF THE ADVERSARY, HOW DO WE IDENTIFY THEM, AND HOW DO WE UNDERSTAND THEIR CULTURE (WHY THEY DO/BELIEVE WHAT THEY DO).**

SITA currently works within the cyber domain and has a working Knowledge of Us and Them as part of the existing framework. The Knowledge of Us contains our capabilities (assets), missions, and vulnerabilities. The Knowledge of Them consists of the adversary's capabilities (assets), capacity, and past behavior. The focus this year is expanding the Knowledge of Them, by developing algorithms which support the analyst in identifying the adversary's intent/goals, and opportunities the adversary might have towards accomplishing their intent/goals. A key challenge will be defining and understanding what are the intent/goals of the adversary, how do we identify them, and how do we understand their culture (why they do/believe what they do). We will also be demonstrating the robustness and scalability of the basic algorithm by applying it to additional domains.

As the NOEM and SITA come together, the combined capability will provide a decision maker with greater awareness into the range of alternate futures of state and non-state actors. This capability will assist them in exploring the realm of the plausible in order to support greater understanding of current conditions and possible outcomes in terms of impact and plausibility. It will lead to more robust, targeted courses of actions as it provides better insight into potential adversary recourses and ramifications of planned blue forces actions (to include unintended consequences). This capability will provide greater variance in training scenarios for blue forces, minimizing surprise, while bounding the realm of the possible, better preparing our forces. It could also afford the user with better utilization of limited intelligence, surveillance, reconnaissance (ISR) resources as they increase our insight into what the adversary can or is currently doing against various blue force assets or missions. Ultimately this will support the fight through. NOEM and SITA are being developed by a team of government engineers (AFRL/RI, in collaboration with AFRL/RH and AFOSR) and contractors (on-site, off-site). The efforts described here are targeted to become part of the future Combat Plans Divisions (Strategy Cell) as part of both the initial planning of friendly courses of action (COAs) and the continuous assessment of mission execution. ♦

# FEATURE ARTICLE ONGOING PHASE I HSCB SMALL BUSINESS INNOVATION RESEARCH (SBIR)

**By John Boiney, PhD**

The Small Business Innovation Research (SBIR) Program provides funds available to small businesses for research that pertain to a Department of Defense (DoD) research interest or need. The Office of the Secretary of Defense (OSD) sponsors general SBIR themes, one of which is Human, Social, Cultural, and Behavioral (HSCB) Technology. Topics under each theme are submitted by government scientists and engineers, down-selected, and a public call for submissions follows. The HSCB Modeling Program Director, CAPT Dylan Schmorrow, is the OSD SBIR Human Systems representative and, as such, part of his role is to coordinate all OSD Human Systems SBIR topics and advocate for the research.

A successful SBIR project may pass through three phases. Following competitive selection, a successful Phase I will typically be one half-person year effort over a period not to exceed six months, with a dollar value up to $100,000. These efforts typically involve feasibility assessment and design specifications, rather than full-blown research and development efforts. Following another round of competitive selections, a successful Phase II will encompass a full-blown effort to develop a prototype based on the analyses done during Phase I.

An important point to note is that to be truly successful, the product of the SBIR must transition to a customer—by having a major program of record agree to assign acquisition dollars to integrate the resultant product into their program, by having a commercial customer agree to contribute additional funds, or perhaps by a warfighter agreeing to implement the capability in their environment. Historically, these negotiations have been delayed until late into Phase II. However, with the start of the FY09 SBIR call (OSD SBIR 9.2), Phase I efforts are required to demonstrate a proposed transition pathway prior to Phase II funding. In April of this year we kicked off a total of twelve of these SBIR Phase I projects from three topic areas. To date, approximately half of these projects have identified one or more possible transition paths, with one already receiving additional funds in Phase I from these customers, and several others in the process of negotiating memorandums of agreement (MOAs) or other, similar forms of transition agreements. In the remainder of this article, we summarize the topics and Phase I awards being coordinated by CDR Joseph Cohn.

## Topic: A Cultural Architecture Generator for Immersion Training in Virtual Environments

Increasingly, military missions focus on small teams, rapid deployment, and global reach. Cultural training is often omitted from deployment preparations. Virtual Environment (VE) training systems offer a potential solution since they have a small footprint, rely on software rather than hardware, and offer scalable training. The objective of this topic is to develop a platform-independent architecture for providing rapid, on-demand and up-to-date cultural immersion training for warfighters who are either operationally deployed or in their pre-deployment workup cycles. Performers were invited to submit proposals to develop a Cultural Architecture Generator (CAG) that would:

**THIS PROJECT WILL EXTEND EXISTING PLUG-AND-PLAY CULTURAL ARCHITECTURE (PNPCA) TO WORK WITH MULTIPLE MODES OF HUMAN INTERACTION (SPEECH, GESTURE).**

- Include tools for scripting scenarios
- Instantiate behaviors through Computer Generated Forces (CGFs)
- Monitor trainee performance
- Provide a transparent, platform independent interface for feature representation within a virtual environment

### Projects

- **Plug and Play Cultural Avatars for Training (SoarTech with Vcom3D)**

  This project will extend existing Plug-and-Play Cultural Architecture (PnPCA) to work with multiple modes of human interaction (speech, gesture). It will also adapt existing auto-data-ingest work (from other projects) to make models and scenarios current. Challenges to be overcome include: a) currently available COTS speech and gesture tools not very reliable, meaning that the system under development must be tolerant of recognition errors, and b)available data for use in driving the CAG is typically not available in a computer-friendly format meaning that the data ingest must be able to "gracefully degrade."

- **C-Core, A Framework for Workflows in Cultural Content, Ontology, and Resource Engineering (Alelo)**

  This project will deliver a framework for workflows yielding authorable, culturally-aware, platform-independent communicative agents. The framework will support data development and formal modeling of historical data management for unit and regression tests. It will use socio-cultural data from a variety of sources: first-person interviews, HUMINT, and published media.

- **Culturally and Socially Aware Believable Agents (Charles River Analytics)**

  This project will develop a flexible, modular agent architecture that provides easy-to-use tools for non-experts to develop agents using real-world updates, combined with

high quality agent interactions to support on demand targeted training and mission rehearsal. This approach includes developing unified, multimodal behavior generation and recognition capabilities, along with Measures of Performance and Measures of Effectiveness.

## MILITARY FORCES THAT CAN ACCURATELY FORECAST HUMAN BEHAVIOR AND USE IT WISELY HAVE AN ADVANTAGE.

### Topic: Algorithmic Behavior Forecasting

Military forces that can accurately forecast human behavior and use it wisely have an advantage. At present, these analyses are performed by humans who bring their own biases, leading to faulty recommendations and inaccurate behavioral forecasting. Moreover, analyses require a significant amount of time to develop, in large part due to the dynamic nature of the source information. Advances in modeling and simulation software have produced increasingly accurate analytic capabilities of human behavior. There are also suggestive advances in methods like genetic algorithmic modeling of human behavior in areas like marketing and advertising industries. The objective of this topic is to develop a tool that will provide an accurate forecast into the cultural and social behaviors of a domestic or foreign target population to enable more accurate and effective decision making.

#### Projects

- **FACETS (Aptima)**

  This project will develop a model-based tool which fuses detailed polling data with timely reports and news to estimate cultural dynamics. As part of Phase I, Aptima will develop a general scenario based on existing polling data. This will represent the 'baseline' into which additional data will be fused – including open-source news and blogs, and extract topics and opinions from the news and blogs. To demonstrate feasibility, they will then inject events and opinions into an agent-based model to demonstrate the integrated approach.

- **Gleaner (Soar Technology)**

  The objective of this project is to support "what if" analyses of proposed military and other types of activities. Soar Tech's approach centers on finding and extracting information about goals and objectives in open sources (newsfeeds, television, radio broadcast, etc.) and using that information to rapidly generate simulation models. In Phase I, Soar Tech will characterize indicators of goal and objective information and enhance their pattern extractor as well as develop a way to inject actions into the simulation model.

- **MAMBA (Edaptive Computing)**

  This project will develop techniques that forecast the behavior of groups and populations. The effort currently focuses on providing intuitive tools that facilitate fast scenario modeling and analysis, extraction and incorporation of cultural variables from archival data, and moving from initial small scale scenarios to field-relevant scenarios.

- **CASCADE (Intelligent Models, Inc)**

  The objective of this project is to automate the assembling of relevant knowledge, the projection of a population's responses, and the subsequent prediction of cascading effects. Intelligent Models will build a hybrid socio-cultural terrain dynamics and behavior simulation system prototype. They will use automated data-mining and sense-making technologies for input filtering and analysis, neuromorphic algorithms for behavior simulation and projection, and layered Sensor Observation Services (SOS) cartography technologies for interactive visual analytics.

- **Forecasting Cultural Analysis and Simulation Tool (SET Corp)**

  This project will use agent-based simulation and socio-cultural modeling to identify direct and indirect effects of diplomatic, information, military, and economic (DIME) actions. The work involves three major segments: extending SET Corp's S-CAT Model to address cultural primitives and courses of action (COA); developing COA manipulation; and extending their Monte Carlo simulation to ensure appropriate substitution of actions and actors.

### Topic: Using Serious Games for Socio-Cultural Scenario Training

The training and cost effectiveness of serious games make them attractive for military applications. However, DoD missions involve a wide range of non-kinetic activities (Phase 0 to Phase 4) in a wide range of socio-cultural environments. Developing "one-off" game-based training for each is cost prohibitive. This topic focuses on developing a low cost, portable Serious Game-based development tool that will quickly train warfighters on a wide range of HSCB knowledge by incorporating near real-time data and putting authoring in the hands of the user.

#### Projects

- **Tools to Create Serious Games with a Cultural Context (Vcom3D, Inc. with Soar Tech)**

  The objective of this project is to develop low cost, portable software tools for quickly training warfighters to operate

CONTINUED ON PAGE 12

# SPOTLIGHT KEITH GREMBAN, PHD

In the spring 2010 issue of this newsletter, we posed four hard research challenges and answered two of them. This summer edition now answers the remaining two research challenges. The first, discusses the need for accurate, usable models that can deliver results in a timely manner. When researchers develop models, they frequently rely on neat, clean datasets that contain exactly the right information and data points required for that particular model. The data that is available to model end-users is typically not curated and is most likely not as complete as a laboratory data set – we need models that are able to ingest "real" data from the field and still deliver reliable output. Dr. Keith Gremban, Vice President of SET Corporation addresses this challenge below.

Dr. Keith Gremban is the Vice President for Command and Control Technologies at SET Corporation, an SAIC company. He has undergraduate and graduate degrees in mathematics from Michigan State University, and a Ph.D. in computer science from Carnegie Mellon University. Dr. Gremban's professional career spans multiple companies, including Martin Marietta, SRI International, and SAIC, where he has led the development of technologies ranging from robotics to numerical computing. The majority of his recent professional experience has been involved with understanding human cognition and behavior, and using technology to complement human capabilities. He leads a team in Denver, CO of artificial intelligence (AI), robotics, and software and systems engineering professionals who are experts in transitioning research results into applications systems.

Dr. Gremban started his career as a researcher in computer vision, where he became fascinated with the problem of understanding human perceptual capabilities and finding means to exploit this understanding in robotic systems. In 1984-1985, while at Martin Marietta Corporation, he was the perception lead for the Defense Advanced Research Projects Agency (DARPA) Autonomous Land Vehicle where his team developed the first computer vision system capable of guiding a robot autonomously on paved roads.

At Carnegie Mellon University, Dr. Gremban initially studied robotics and computer vision, while pursuing his interests in human perception and cognition. His dissertation extended and applied graph theory, an area of particular relevance to his work in human cognition and human behavior modeling.

Dr. Gremban's interest in human behavior continues to be inspired by human interactions with command and control (C2) systems. Much of his early work involved simplifying the interaction between C2 systems and dismounted warfighters, where the user is cognitively constrained by many demands other than paying attention to his C2 system. The warfighter's focus of attention must be on his surroundings, and the C2 system should not be a distraction. The C2 system must complement the warfighters activities, rather than conflict with them. This philosophy of complementary operation led to the design of a unique hands-free, eyes-free C2 system for human-robot collaboration. The system enables a robot to cooperate with a human team by observing the behavior of the human teammates, and then selecting a complementary behavior.

Currently, Dr. Gremban and his team are applying lessons learned from years of designing behaviorally adaptive C2 systems to the challenges of the HSCB Program. Dr. Gremban is the program manager for the HSCB-sponsored Socio-Cultural Analysis Tool (S-CAT), a system being developed by a team led by SET, consisting of SRI International, SAIC, consulting field anthropologists, and military subject matter experts with experience in operational planning, who are influencing the development of S-CAT to meet the needs of military users. S-CAT is a tool for modeling societies and forecasting the potential consequences of socio-cultural courses of action. It employs two mechanisms for forecasting: inferential reasoning for low-order direct and indirect effects, and agent-based simulation for long-term consequences. S-CAT provides a socio-cultural framework and representation that is accessible to military planners, analysts, social scientists, and computer scientists and bridges the communications gaps between them.

Several observations drove the development of S-CAT. First, Dr. Gremban and his team realized that modeling must be a continuously ongoing activity. The very act of modeling influences and changes the target society, which also continuously changes from natural causes. As a result, the S-CAT knowledge base is designed to continuously evolve. The concept includes a preliminary model of basic knowledge that one can use to develop high-level forecasts, and an interface that enable users to continuously add observations as behavioral rules in the model. The tool can generate forecasts and explanations that use the "wisdom of the crowds" to consider each rule as another voice in the crowd. Each voice provides useful information for the operational analyst, planner, or decision maker.

Forecasting can never be exact. Human behavior is dependent on too many variables that cannot be anticipated or modeled. Hence, S-CAT focuses on forecasting plausible outcomes. The goal is to stretch the mind of the decision maker by presenting the possible consequences and providing explanations linking actions to effects.

Socio-cultural analysis is an increasingly important capability for the modern military, as well as other governmental and private organizations. Success in the social, economic, or military battlefields depends on an understanding of what makes people tick. Yet, it is impractical to require every military or civilian leader to be a cultural anthropologist. S-CAT is designed to enable users to understand the possibilities of human socio-cultural behavior without requiring extensive and specialized training. ♦

The second research challenge stated that multiple approaches to HSCB-oriented training exist, including, but not limited to: in-person, web-based, avatar-based, game-based, culture general, and culture specific. Determining which to choose for particular missions, learning outcomes, and/or personnel is an ongoing challenge. Dr. Allison Abbe, of the Army Research Institute, discusses the HSCB Program's efforts to address this challenge in the spotlight below.

The HSCB Program and other defense programs have funded research and development on cultural training for military personnel. Current instructional programs and training products target cultural learning objectives in a number of ways, including distance learning courses, live simulations, virtual simulation, and smart cards. Using these various interventions at different points in the career and deployment cycle is consistent with the career development approach to language and cultural capabilities that the Services have adopted. However, some training interventions may have more impact than others, and some will require more resources to sustain than others. Given constraints on resources and training schedules, training interventions must be planned and selected to maximize efficiency. How can we build a program for developing cultural capability that progressively builds the skills and knowledge needed? At what points in the career or deployment cycle will cultural training have the most impact? Where can cultural learning be integrated with other training objectives to reduce the training load?

Within the U.S. military services and DoD, there has been emerging convergence on a general common framework for the consideration of culture-related capabilities, identifying three broad learning domains – regional knowledge, foreign language, and cross-cultural competence. Whereas regional knowledge and foreign language are specific to the people of a particular geographic locale, cross-cultural competence is the set of knowledge, skills, abilities, and attitudes (KSAAs) that enable an individual to function within and learn about a culture without *a priori* knowledge of that particular culture. It may also provide the general foundation that would enable more rapid acquisition of specific regional knowledge in mission-focused training. These learning domains have traditionally been viewed as relevant primarily for specialist personnel, such as interpreters and foreign area officers. Developing and assessing these skills in general-purpose forces and at lower levels of proficiency than that of specialists is a significant challenge.

ARI is conducting a three-part project focusing on broad challenges for training and development and on assessment of cultural capabilities. The first objective is to identify training and development methods that best address cultural KSAAs. The output will be a set of conclusions and recommendations regarding the expected developmental sequences of the KSAAs, highlighting particular forms of interventions that aid in the development of these KSAAs, potentially to include training, education, professional assignments, self-development, and other experiences.

The second objective is to identify promising measurement methods for assessing those KSAAs for purposes of training evaluation, training needs analysis, and potentially for selection and assignment. The output of this process will be a set of conclusions and recommendations regarding appropriate assessment methods for the KSAAs, assessment methods that can be incorporated into training and education programs, and any research gaps that need to be addressed in order to develop assessment tools for cross-cultural competence. For each of these objectives, academic researchers will provide advisory input. We are working with psychologists, anthropologists, and psychometricians whose expertise spans international education, expatriate management, employee development, practical intelligence, and testing and assessment.

The third objective is to identify mission and other contextual characteristics that influence the nature and amount of cultural capability required for different roles. Beyond a basic level of cultural awareness, personnel who deploy will likely have performance requirements related to culture that could differ dramatically. Missions and roles vary in terms of the type of knowledge and skills required, as well as the overall relevance of cultural differences. This research will produce a set of critical incidents that can be used in training development and a taxonomy of mission and situation variables to help determine what cultural knowledge and skills should be targeted in training for certain missions or roles.

This research will provide foundational input to developing and sustaining programs and tools that enhance the Servicemember's ability to perform in the socio-cultural aspects of his mission. The goal is to provide input to training developers, force development programs, and other decision makers on methods for developing and assessing cultural capabilities. Findings will help identify training gaps and improve cultural training development, providing conceptual models for learning and methods for evaluating training effectiveness. This research will also help advance the HSCB Program by providing foundational knowledge for integration of user or trainee capability into computational models. These models of operator cognition are a critical consideration in training and performance support tools (Foster & Fletcher, 2003). For example, one likely application is the development of learner models for adaptive training and simulation. Methods to determine the learner's state of intercultural development are currently lacking for such individualized instruction. Thus, this research will better enable the design of training, education, and simulations, as well as provide methods for determining which of these interventions are more and less effective. ♦

## Reference

Foster, R., and Fletcher, D. (2003). Modeling the user for education, training, and performance aiding. Paper presented at the NATO Human Factors and Medicine Symposium, Advanced Technologies for Military Training, Genoa, Italy.

in an unfamiliar cultural environment using immersive serious games. It will extend the team's proven Plug-and-Play Cultural Avatar Architecture (PnPCA) to integrate new human interface modalities, and to demonstrate reuse across a wide range of platforms. The team expects to develop partnerships to meet critical technical challenges.

- **FRED/JACK (CHI Systems)**

  With their FRED/JACK project, CHI Systems expects to deliver low-cost game-based HSCB training centered on light-weight intelligent, cultural avatars with multimodal interaction. A major technical challenge is designing and programming avatars that can recognize and react to users' facial expression, body language and gestures.

- **C-GAME, Game-Based Cultural Competency Training (Alelo)**

  To deliver this training tool, Alelo will leverage and extend its VRP agent architecture and situated culture ontology. It will augment the game environment with cultural annotations and intelligent scaffolding. Assessment of the training efficacy will be based on reactions of agents to learner actions.

- **Authoring By Cultural Demonstration (Aptima, Inc)**

  The objective of this effort is to design and demonstrate a domain and platform independent tool to author, generalize, and execute/monitor socio-cultural training. To accomplish this objective, Aptima will develop a use case and templates of gestures, behaviors, and dialog to focus tool development. The performer expects to demonstrate the feasibility of a low cost, culturally-enabled scenario authoring and training tool.

## Conclusion

The measure of Phase I success includes technical performance toward the topic objectives and evaluations of the extent to which Phase II results would have the potential to yield a product or process of continuing importance to DoD and the private sector. Thus, awardees are strongly encouraged to identify and begin working from prospective transition targets as soon as possible in Phase I. To that end, awardees have access to technical and programmatic support from CDR Cohn and others who support the SBIR and HSCB Modeling Programs. ♦

## CALENDAR OF UPCOMING CONFERENCES AND WORKSHOPS

| Date | Event | Location | Sponsor | Website |
|---|---|---|---|---|
| **July 17–20, 2010** | 2010 AHFE International<br>3rd International Conference on Applied Human Factors and Ergonomics<br>*Jointly with*<br>1st International Conference on Human Factors and Ergonomics in Healthcare<br>1st International Conference on Cross-Cultural Decision Making<br>13th International Conference on Human Aspects of Advanced Manufacturing | Miami, FL | | **www.ahfe2010.org** |
| **August 16–17, 2010** | Unifying Social Frameworks: A Workshop | Washington, DC | National Academies of Science | |
| **September 27–October 1, 2010** | 54th Annual Meeting of the Human Factors and Ergonomics Society, 1st International Conference on Cross-Cultural Decision Making | San Francisco, CA | Human Factors and Ergonomics Society | **www.hfes.org/web/HFESmeetings/2010annualmeeting.html** |