OPERATING WITHIN THE HUMAN DIMENSION:

COLLECTION AND ANALYSIS CHALLENGE

WORKSHOP REPORT

24-25 April 2008

PREPARED FOR:

UNITED STATES DEPARTMENT OF DEFENSE

Joint Staff/J3
United States Strategic Command/Global Innovation and Strategy Center
Director, Defense Research and Engineering
Office of the Under Secretary of Defense, Intelligence
Office of Advanced Systems and Concepts

UNITED STATES DEPARTMENT OF HOMELAND SECURITY

University of Maryland,
National Consortium for the Study of Terrorism and Responses to Terrorism

PREPARED BY:

SARAH CANNA
NATIONAL SECURITY INNOVATIONS (NSI), INC.
301-466-2265
scanna@natlsec.com

ROBERT POPP, PHD
NATIONAL SECURITY INNOVATIONS (NSI), INC.
781-864-1347
rpopp@natlsec.com

STEERING GROUP MEMBERS

Workshop Coordinators

Robert Popp	National Security Innovations, Inc.
Sue Numrich	Institute for Defense Analyses

Steering Group Members

Kelcy Allwein	Defense Intelligence Agency
Frank Connors	Defense Threat Reduction Agency
Bert Davis	Army Engineer Resource and Development Center
Laurie Fenstermacher	Air Force Research Laboratory
Carl Hunt	Institute for Defense Analyses
Kathleen Kiernan	The Kiernan Group
Brian Meadows	Space and Naval Warfare Systems Center
Patti Morrissey	Office of the Under Secretary of Defense, Intelligence
Jennifer O'Connor	Department of Homeland Security
Lt Col Mike Shannahan	Strategic Command, Global Innovation and Strategy Center
Kathie Smarick	Department of Homeland Security, START Program
Allison Smith	Department of Homeland Security
Ann Speed	Sandia National Laboratories
CAPT Todd Veazie	Joint Staff

OPERATING WITHIN THE HUMAN DIMENSION: COLLECTION AND ANALYSIS CHALLENGE

EXECUTIVE SUMMARY

One of the major challenges the US has faced during the opening decades of the 21st Century is the growth in the number and international reach of insurgents, terrorists, WMD proliferators and other sorts of nefarious groups. Among many others, the current situation in Iraq illustrates the difficulty of countering insurgencies once they have become violent and operationally well-established. The apparent transition of contemporary war away from the traditional physical battlefield towards the human element as the battlefield poses new challenges for both intelligence and operations – collections and analysis within the human domain requires new thinking and new methods.

The purpose of this workshop was to identify approaches to recognize and anticipate the conditions in which insurgents, terrorists, WMD proliferators and other sorts of nefarious groups tend to emerge, become operationally established, and choose to employ tactics counter to the interests of the US national security objectives. An underlying presumption of this workshop is that tackling these movements during their emergent phases is less costly in resources and lives, with more non-lethal options available for consideration, than after they have become fully established. Novel approaches to collections and analysis of the interactions within and between elements of the "human dimension" may lead us to greater potential for proactive rather than reactive capabilities, and certainly non-kinetic capabilities, not only within the military context, but within all components of national power.

Clearly, focusing on emergent violent groups and their activities will necessitate a shift in how the nation and its allies collects, analyzes, and acts upon information from and about these volatile environments. The workshop brought together a diverse group of physical and social scientists, academics, analysts, intelligence collection managers, operational planners, and decision makers to discuss novel techniques and methods for a more proactive approach to address the human dimension associated with insurgents, terrorists, WMD proliferators and other sorts of nefarious groups, and the impact of such an approach on training, planning, tasking, collection, analysis, and decision making in the coming decades.

EVOLUTION OF VIOLENT EXTREMISM

There is an important distinction between activism and radicalism. Activism is legal and nonviolent political action, while radicalism is illegal and violent political action. Activism is not necessarily a conveyor belt to radicalism. Individuals can be radicalized from any layer of society including neutral parties, sympathizers, and activists.

Government Sponsored Research on Violent Extremism

The Department of Homeland Security (DHS) is is laying the groundwork to conduct in-depth case studies of individuals engaged in terrorist violence in the US. At the group level, DHS is studying terrorist group rhetoric with a focus on both groups that have engaged in terrorism and groups that would be considered radical but that have not engaged in terrorism. The study will look at groups with similar ideologies and goals and try to understand why some groups choose violence and others do not. DHS will also work to identify characteristics of groups who participate in IED attacks. At the community level, DHS is involved in two primary activities. The first is evaluating whether survey data can shed light on radical activity in the US. Finally, DHS is working with START to study the effectiveness of counter radicalization programs conducted in five countries including Yemen, Indonesia, Columbia, Northern Ireland, and Saudi Arabia.

The Behavioral Analysis Unit is the newest organization within the Federal Bureau of Investigation. It is an operational unit that conducts research. Counterterrorism research within the unit is still in its infancy stages;

however, some projects are currently underway dealing primarily with operational threat assessments. The most pertinent research efforts are in collaboration with DoD intelligence units. This work primarily focuses on radicalization by looking at the radicalization process in historic case studies. The FBI is also involved in another research effort in coordination with the WMD directorate, to look at 50 domestic cases in which offenders used WMD in the US.

The Defense Intelligence Socio-Cultural Dynamics Working Group (SCDWG) was created in 2002 to develop an enterprise solution to institutionalize Socio-Cultural Dynamics intelligence analysis. The group was asked to address the full scope of issues, from policy to tactical. The group was asked to define the role of intelligence in this area as well as to identify the "right tool for the right job."

The Marine Corps Intelligence Activity (MCIA) is working on the next generation of cultural intelligence. They are working on visualizing cultural intelligence by "putting culture on the map." Their second project is to support influence operations and irregular warfare by building partner capacities focused on the values and themes of sub-national groups. The third is cultural vignettes, which are short and focused products designed to create a specific cognitive map for each mission set. The fourth is building an analytic capacity by supporting uniquely trained cultural analysts, ensuring 24/7 reachback for deployed unites, and creating deployable analytic teams for contingency operations.

UNDERSTANDING SOCIO-CULTURAL DYNAMICS: TRADECRAFT AND OBSERVATIONS FROM A DISTANCE

Studies to understand a country and its people must be planned and carried out with nuance. Ethnic maps, a critical tool for US forces, have long been a political issue. It affects resource allocation, boundaries, and political formulae. In countries experiencing ethnic conflict, building a credible demographic baseline must be done before lasting geopolitical stability can be realized. Maps can help researchers and analysts understand the cultural geography of conflict. Chorological analysis, relating to the description or mapping of a region, can reveal causal linkages and patterns of political and social behavior that researchers were previously unaware of. Mapping the human terrain is needed to understand these complex problems.

Culture at a distance methodology from 1942 mostly comprised of several sources. Researchers studied travelers' accounts, ethnographies, histories and other second hand sources including social science works. Content analysis finds its roots in these attempts. Post World War II, the DoD funded a great deal of social science research although only a small part of it was "culture at a distance". They focused instead on models such as game theory. Their main interest was in political instability: its causes and indicators.

Today, the DoD has stated that socio-cultural data collection is everyone's problem. However, the DoD does not currently have the knowledge and understanding of non-Western cultures and societies necessary to execute these missions. Because of this, there are several new culture at a distance methods. Computational modeling has become the sine qua non of culture at a distance method, but it suffers from considerable handicaps. Its emphasis on computational engineering leads to "everything looks like a nail" thinking. The problem with models is that many do not scale gracefully from explaining small scale issues to larger scale ones. Models are culturally and socially ignorant, which is frequently the case when engineers develop "new cultural theory" with little or no training in the fields of social science.

Socio-Cultural Tradecraft and Open Source Collection Requirements Management

Emerging violent non-state actors use or share a common story about their involvement to motivate and empower their groups, and attract and mobilize their audiences. The degree by which infringements are interpreted, personified, and perceived is the variable to which violence emerges. It is typically a conflict between the haves and the have not's over a particular issue. Modeling these relationships requires a four prong approach: anthropological research; open source collection; human intelligence collection; and environmental research. The model will synthesize that information to assess behavior.

Cultural geography is the study of spatial variations among cultural groups and the spatial functioning of society. Urban geography is the study of how cities function, their internal systems and structure, and the

external influences on them and the study of the variation among cities and their internal and external relationships. The human geography areas of interest include economics, demographics, politics, culture, combined social implications, and technology/infrastructure. Cultural geographers face several challenges. First, experiences in Afghanistan and Iraq have shown that neighborhood level data is necessary for operational and tactical planning (Intelligence Preparation of Battlefield). Second, there is a lack of reliable data sets. For example, there has been no census in Lebanon since 1932. Third, closed countries, such as North Korea, pose a problem. Fourth, no city level data exists or it is difficult to obtain in grey literature. Fifth, there is a lack of current academic studies. Therefore, researchers must take a multidiscipline approach and use sources from a variety of fields to complete their research.

Remote Observing Using Large Data Sets

The nexus for computer forensics is addressing large data sets. Researchers and developers are working with open source information and applying analytic tools to it. They have discovered that it is considered novel in the US to use the same tool to address open source and classified datasets and merge them. One large data set, INSPIRE, was developed to allow an analyst to look at a large number of text documents and look for words and phrases. The program graphically portrays the output. A subsequent project is porting INSPIRE over to high power computing environment.

Content analyzers are another tool for working with very large datasets. The tool makes coding schemes in the process of creating large data sets from large text sources. It has an automatic generation process that is transparent. From this data, the analyst produces dashboard like metrics to describe the data. This has created an impetus to create measures of effectiveness and validation. The team discovered that people and computer coders make different types of errors. Humans make boredom errors. Their accuracy always goes down over time. However, if a computer sees a new word, it will not code it. The operation code measures how you present self and others in text on cooperative or conflictive dimensions. The tool can also extract and compare entities for distrustfulness. They use a combination of tools to look for themes and people. It is a pragmatic way to build structure while reducing the analysts' reading load.

UNDERSTANDING THE HUMAN TERRAIN: TRADECRAFT AND STREETCRAFT METHODS

Applying Law Enforcement Methods to Gathering Human Terrain

An appropriate balance of skills, knowledge, and abilities need to be incorporated into an analysis team. It needs to be half Army Ranger company and half artist colony. Good analysts are adaptable, intellectually flexible, and have a high degree of comfort with ambiguity. An analyst's job is simple: they read, write, and present. In the development of a good analyst, no preparation matches analyst's comfort level with ambiguity. Furthermore, the depth and breadth of an analyst's rolodex is more important than personal understanding of a region or topic. Analysts must also exhibit finesse - the practice of combining knowledge, thought, and intellectual stamina. Personal conviction of paying attention to a few small details that have enormous impact on outcomes or intelligence is critical. Good analysts are those who care as much about operational environment as producing a final product.

Approximately 90 percent of the food we eat, clothes we wear, and products we use are imported daily via commercial maritime transport. Threats to ships and ports include drug smuggling, stowaways, and terrorists. Drugs are often creatively hidden in cargo, such as cocaine stashed inside individual bananas. Drug couriers can also be hidden away in cargo shipments. These could easily be terrorists or terrorist materials. According to a US Coast Guard report, 25 Islamic extremists entered the US onboard commercial cargo vessels. Narco traffickers are more frequently taking consignments of extremists to the US. There is an increasing level of cooperation between MS-13, narco traffickers, and Islamists. The human collection component is critical to stopping these illegal activities because Customs cannot provide the level of analysis needed.

The Department of Commerce's task is to prevent the export of US dual-use goods and technology that may be used by rogue states or terrorists to make chemical, biological or nuclear weapons. The Office of Export Enforcement's priorities are WMD proliferation, terrorism/terrorist support, and unauthorized military/

government use. Commerce is unique in that it has civil penalties, in addition to criminal penalties, in which one only needs the preponderance of the evidence to convict. The challenge of identifying and stopping illegal dual use exports is making sure at every step of the process that all of the people involved can validate the end use. It gets murky because end users lie and some components are working with the end users. Globalization is also a national security challenge and many transactions are moving over the internet. The US government does not move at the speed of business.

The Criminal Investigation Task Force (CITF) was created in early 2002 by the DoD to conduct investigations of detainees captured in the Global War on Terrorism. CIFT became the investigative arm of the DoD after OSD was given the mission to prosecute the detainees. OSD transferred the responsibility to the Army, where the CITF now falls under Army Criminal Investigation Command (CID). The need for analysts and intelligence professionals was huge, especially for law enforcement officials not used to working with classified materials. Likewise, law enforcement people need to communicate with the combatant commanders. Terrorists groups often act like organized crime groups. The intelligence community may not have much experience in this, but law enforcement does. It also has experience in conducting interrogations. To conduct interviews effectively the interrogator must put all biases aside and just accept the culture for what it is. The interrogator must get down to the bedrock of cultural foundation.

Applying Social Science Methods to Understanding the Human Terrain

Social science is the application of consistent, rigorous methods of research and analysis to describe or explain social life. Its purpose is inference: using facts we do know (data, observations) to learn about facts we do not know (theories, hypotheses). It is used to make descriptive inferences (semantic program); to place or make observations within a conceptual framework to allow people to understand a phenomenon; and to make causal inferences (syntactic / pragmatic program). Because of the large number of potential causal factors for any social phenomenon, establishing causation is difficult. If we can establish the cause of a social phenomenon, then we can make policy to affect that phenomenon.

Data collection involves various methodologies such as ethnography: participant observation; surveys: interview processes (formal, informal, group, individual, structured, unstructured); record or document review; and history: conducting interviews or reviewing documents about the past. Data analysis involves direct interpretation: analysis by an individual's reflection and synthesis; quantitative analysis: using standard methods of statistics to ascertain relationships; and formal modeling: analysis by creating a formal system that mimics the world. In evaluating which method is most appropriate, it is important to keep in mind that multidisciplinary methods reinforce one another.

Field research in Iraq is particularly complicated. Researchers must use semantic programs and descriptive inference as well as create a pragmatic program with causal inference. Some of the challenges include sampling, reactive bias, and interpretive bias. Coping with these problems is done by working through local researchers, geographic sampling and purposive sampling, redundancy of methods, and redundancy of research networks. On the ground analysis is informed by area studies, topic specialization, and occasional forays into population.

There is a great need to improve intelligence support to stability operations. To do this, analysts must embrace the importance of understanding the human terrain. There are three key elements to the human terrain. The first is developing a socio-cultural dynamics data network and repository. The second is developing data gathering, visualization, and analytic tools. The third is recruiting, training, and deploying experts to support decision-making.

To understand counterinsurgency, you must first understand and define the elements involved. Collecting information for COIN operations involves identifying key groups in a society and representing where they are located on a map. For each group, researchers seek to identify security (level, sources, threats); income and services (level, sources, gaps); beliefs and communications systems (narratives, symbols, norms and sanctions); and authority structures & figures (identity, structures, levels of authority).

Soldiers as Human Terrain Sensors

The Civil Affairs (CA) information objectives were to provide multi-dimensional situational understanding, provide situational analysis, and to work on the development of possible solution sets. The foci of CA operations are human terrain identification and prioritization (sphere of influence management) and civil reconnaissance. Within the human terrain network, there are sphere of influence engagements. These works identified political/tribal leaders throughout the district and their areas of influence to facilitate reconciliation efforts. They also identified facility managers/public works workers to facilitate transition to Iraqis fixing Iraqi problems. They also prioritized efforts to optimize limited resources for engagement.

There are three critical nodes in the CA analysis process. The first is collection. The information is the critical base element to the process. Without this, there is no system. The second is consolidation. Analysis and understanding the information makes the process function. It is an art and a science. The third is dissemination. Sharing the situational understanding is the key to success. Dissemination must flow up and down in order to coordinate and integrate with partners.

Provincial Reconstruction Teams (PRTs) have become a key tool for the international community to assist Afghanistan in becoming a secure and self-sustaining Republic. They represent, at the local level, the combined will of the international community to help the government and the civil society of Afghanistan become more stable and prosperous. PRTs, due to their provincial focus and civil-military resources, have wide latitude to accomplish their mission of extending the authority of the Government by improving security, supporting good governance and enabling economic development. This engagement of diplomatic, military and economic power by Nations at the provincial level allied with the wide latitude to accomplish their mission has been a strength, as it provides flexibility of approach and resources to support the provincial government structures and improve security.

The Humanitarian Information Unit (HIU) serves as a United States Government (USG) interagency center to identify, collect, analyze and disseminate unclassified information critical to USG decision makers and partners in preparation for and response to humanitarian emergencies worldwide, and to promote best practices for humanitarian information management.

Civilians as Human Terrain Sensors

In order to operate effectively in Afghanistan, warfighters need to understand the cultural context and history of the nation. An anthropology background by itself is not always useful. Human Terrain Teams (HTTs) and others need culture specific knowledge to achieve their mission in foreign countries. An open source portal exists to provide a civilian based source of socio-cultural information on Afghanistan. The portal facilitates a subject matter expert network to encourage interaction among the key players including the Department of State (DoS), non-governmental organizations (NGOs), the US Agency for International Development (USAID), Brigade Combat Teams (BCTs), and soldiers. The site drills down to the provincial level where 21 of 34 provinces have been completed. The site provides information on detailed maps, refugees, education, health, topographic information, etc.

SUPPORTING ANALYTIC DISCIPLINES

Modeling and Visualization Workshop Summary

The objective of the Social Science Modeling and Visualization Workshop, held in January 2008, was to provide a "mixing bowl" for social scientists, modelers, researchers and government stakeholders to discuss the state-of-the-art in methods/models/visualization and their potential application in SMA efforts. Models can clarify complex situations, test assumptions, aid decision making, explore co-evolutionary motivations and bound the expected and outlier behaviors. Modeling can capture dynamical, non-linear processes. They can be multi-level, multi-scale behavior representations. Other benefits include that models can assist in training decision makers on the consequences of their policies/decisions. Visualization is a critical component of modeling. Therefore, it needs to be considered from the beginning of a project. Visualization serves a variety of functions include support of knowledge management, contextual visualization, relational

visualization, identification of key features/nodes/information, understanding of decisions/complex processes, and evaluation/inspection of the data. Models must be verified, valid, and credible. The theoretic framework of social science modeling and visualization still needs work. Opportunities and challenges include the need for better, scientifically grounded theories. The federal government is interested in social science, but struggling to determine how to support it. It will take 4-5 years to reach robust funding levels.

Evolutionary Agent-based Modeling and Game Theoretic Simulations

Agent Based Modeling (ABM) is successful in the social science arena because as in the traditional approach to formal modeling, the goal is reducing the social landscape to a set of meaningful variables and specific mathematical equations. ABM brings to social science the ability to begin with a social landscape of entities in the world. It can identify basic relationships between actors. All models are developed or designed to answer a class of questions. ABMs are typically designed to answer questions that are intractable through earlier and more traditional mathematical or statistical methods, such as complex emergent patterns in large systems of agents. The beauty of ABM is that it can conduct experiments about shifting policies. All this can be done in silicon because you cannot manipulate players in the real world. ABM is a specific kind of simulation. It provides an experimental setting for finding patterns that are not very obvious.

Game theory is a branch of applied mathematics that formally models strategic behavior. Game theory is an approach to aid in articulating and understanding important factors of conflicts / disputes / coalitions. Game theory is neither prescriptive nor descriptive; nor is it normative. It is a theoretical tool. Game theory does not provide an answer; it provides a helpful way to frame the question to better understand the situation. If the assumptions are correct, game theory will provide an answer. Some assumptions of the theory include (bounded) rationality, complete information, and preference hierarchy for players (utility). In conclusion, it is important to remember that people are not agents. Game theory requires assumptions about behavior. It provides insight and understanding, but people are more complex than mathematical models. Finally, as in any model, if you put garbage in, you get garbage out.

Data Fusion / Integration and Detecting Patterns in Heterogeneous Data Sets

The defense and intelligence communities are making great progress in the social sciences. Many agencies now realize the importance of social science approaches, in combination with more traditional approaches. Often, people want to know why if we can go to the moon we cannot accurately model social science. It is because people are complex intelligent agents with feedback. They have multiple incentives, multiple allegiances, and multiple groups. Humans also exhibit adversarial behavior. They provide inaccurate and misleading information. Behavior is not invariant. There is a hierarchy of challenges in data fusion. The first is registration and entity resolution. The second is group and network detection. The third is feature construction and recognition. The fourth is complex event detection. The fifth is that concepts in models are inherently fuzzy. There is a prediction versus risk relation. And risk brings additional scrutiny and preparation challenges.

Data fusion technologies encompass a variety of characteristics. It is historically a deductively-based inferencing / estimation process. The resulting observations and estimates produce random variables optimized for minimum uncertainty. The process is adaptive in various ways. It enables observation management, process adaptation, and beneficial actions. Data fusion models can be used to analyze a broad set of areas. The physical domain is the easiest to accomplish fusion. This includes weapon systems and weather reports. The information domain is harder. This includes symbolic representation and interpretations of data and models. The cognitive domain includes beliefs, values, and emotions and is the most difficult to fuse.

Efforts are underway to create an automated multi-source, multi-sensor, multi-data fusion tool. One way they are doing this is by looking at textual information. They take documents of different types, use ABA tools and extract antecedents. The tools can grab information, bring it together, and end up with predictive models. Working toward multi-focused analysis, the tool must overcome the problem of sorting through tons of information.

Large Data Sets and Knowledge Extraction Techniques

Large data and visualization efforts are becoming an ever growing avalanche within the social science community. Loading all the data is hopeless. Therefore, efforts have focused on trying to load the structure of the data. This results in maximum expressiveness with minimum size. The structure of the data for documents includes terms, entities, and concepts. The structure for transaction, travel, and communication records include people, places, organizations, and relationships. However, smart loading still requires human guidance. A sandbox is an environment for "what if" questions and for constructing arguments. It should have quick access to exploratory tools and answer the question, "How are these actors connected?" The system's job is to track workflow and allow for annotations and assumptions. The analyst can backtrack and replay, branch, and recombine. He may also retract or change assumptions. The output should embody a complete chain of reasoning.

Work is being undertaken to determine whether researchers could construct a scaffold for the data that can say something about the dataset. The geometry in the data set might help pull out information. By comparison if you look at physics, the math needed to model it is very tame math. In biology, the math needed is fundamentally different. Even though researchers have not discovered the laws of social science, there are some regularities, although they have not been captured in math to a sufficient depth.

The scale and speed of the data means that the classic, craftsman approach to machine learning is no longer possible. The answer is a commodity, hands-on model. This is a challenging task. Not only does commodity mean that researchers have to develop algorithms that are robust in the face of noise, skew, and the host of other ills; it also means that they have to work out how to set all algorithm parameters from the data or the context, so that the human analyst need not concern themselves with them. There is no single algorithmic magic bullet. There are lots of pattern recognition algorithms; some are well known, including decision trees, support vector machines, neural nets, nearest neighbors, naive Bayes. Ensembles are the first and the most powerful of the techniques to know about, as they permit you to squeeze all possible accuracy out of your data.

TABLE OF CONTENTS

GLOSSARY OF TERMS

ABM	Agent Based Model
BCT	Brigade Combat Team
CA	Civil Affairs
CID	Army Criminal Investigation Command
CITF	Criminal Investigation Task Force
COIN	Counterinsurgency
CONUS	Continental United States
DHS	Department of Homeland Security
DIA	Defense Intelligence Agency
DoD	Department of Defense
DoS	Department of State
FBI	Federal Bureau of Investigation
GIS	Geographic Information System
GTB	Global Terrorism Database
GWOT	Global War on Terrorism
HIU	Humanitarian Information Unit
HTS	Human Terrain System
HTTs	Human Terrain Teams
HUMINT	Human Intelligence
IC	Intelligence Community
ICEWS	Integrated Crisis Early Warning System
IED	Improvised Explosive Device
IGO	International Government Organization
JI	Jemaah Islamiyyah
MAR	Minorities at Risk
MAROB	Minorities at Risk Organizational Behavior
MCIA	Marine Corps Intelligence Activity
NGO	Non-Governmental Organization
OCONUS	Outside the Continental United States
ODNI	Office of the Director of National Intelligence
ONR	Office of Naval Research
OSD	Office of the Secretary of Defense
PNNL	Pacific Northwest National Laboratory
PRTs	Provincial Reconstruction Teams
SCD	Socio Cultural Dynamics
SCDWG	Defense Intelligence Socio-Cultural Dynamics Working Group
SME	Subject Matter Expert
SSTR	Stability, Security, Transition, and Reconstruction
START	The National Consortium for the Study of Terrorism and Responses to Terrorism
USAID	United States Agency for International Development
USDI	Under Secretary of Defense for Intelligence
USG	United States Government
WMD	Weapon of Mass Destruction

INTRODUCTION

The *Operating within the Human Dimension: Collection and Analysis Challenge Workshop* was convened in response to a request from the Commander, US Special Operations Command in coordination with the US Strategic Command to Lieutenant General Carter F. Ham, Director of Operations, J-3, Joint Staff for assistance in developing a national-level capability to identify and anticipate the action of Violent Non-State Actors (VNSA). The VNSA's intent and capability to acquire, build, store, employ, deploy, or supply weapons of mass destruction (WMD) is of particular of interest. In support of this effort, a multi-agency team sponsored a workshop on 24-25 April 2008, hosted at Directed Technologies Incorporated (DTI) facility in Arlington, Virginia.

The purpose of the workshop is to identify approaches to recognize and anticipate the conditions in which insurgents, terrorists, WMD proliferators and other similar groups tend to emerge, become operationally established, and choose to employ tactics counter to the interest of US national security objectives. An underlying presumption of the workshop is that countering these movements during their emergent phases is less costly in resources and lives than after they have become fully established. Novel approaches to collection and analysis of the interactions within and among elements of the "human dimension" may lead us to greater potential for proactive rather than reactive capabilities, not only within the military context, but within all components of national power.

Clearly, focusing on emergent violent groups and their activities will necessitate a shift in how the nation and its allies collects, analyzes, and acts upon information from and about these volatile environments. The workshop will bring together a diverse group of physical and social scientists, academics, analysts, intelligence collection managers, operational planners, and decision makers to discuss techniques and methods for a more proactive approach of addressing the human dimension in the coming decades.

The focus of this workshop is trying to understand how to observe populations of interest. The first session focuses on how observations are made from a theoretical perspective. The afternoon sessions focuses on putting theory into practice. The difficulty is to learn about populations from a distance (both time and space). The goal is to leave the population undisturbed and to conduct the research and analysis where you have the tools to do so. Day two's session will focus specifically on observing populations from a distance. In the final session, the panels will review how the community is using new technologies to accomplish these processes.

PRESENTATIONS DAY ONE

SESSION 1: THE EVOLUTION OF VIOLENT EXTREMISM

Allison Smith, Department of Homeland Security (Moderator)
Clark McCauley, Bryn Mawr College, START
Todd Helmus, RAND
Amy Pate, University of Maryland, START

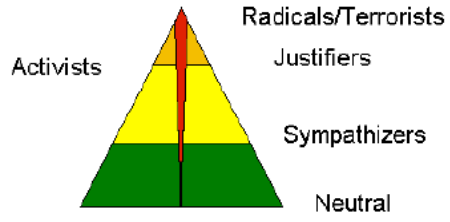
CLARK MCCAULEY

Professor of Psychology and Co-Director, Solomon Asch Center for Study of Ethno-Political Conflict, Bryn Mawr College; Co-Director, National Consortium for Study of Terrorism and Responses to Terrorism (START); Editor, Dynamics of Asymmetric Conflict

Clark McCauley, a social scientist at Bryn Mawr College, spoke about the models and mechanisms of political radicalization and terrorism. He has a particular interest in attempting to understand and measure radicalization.

Political radicalization is denoted by a change in beliefs, feelings, and actions toward increased support of intergroup conflict. The United States after 9/11 provides one example of this. Meanwhile, anti-U.S. radicalization is further defined by the change in beliefs, feelings, and actions toward increased sympathy and support for groups fighting the U.S.

There is an important distinction between activism and radicalism. Activism is legal and nonviolent political action while radicalism is illegal and violent political action.



The image on the left depicts the radicalization process. The pyramid consists of three horizontal layers. The bottom of the pyramid represents the neutral part of the population. The middle layer represents sympathizers. The top layer represents the justifiers. Radicals and terrorists are the red inverted cone that cuts through all of the pyramid's layers representing an individual's ability to become radicalized at any layer, without proceeding through the pyramid layers. Activists are listed on the left because while many people see activism as a step toward radicalization; it is not a necessary stepping stone. Some activists become radicalized, while many do not.

Polling data may be used to understand the distinctions between the layers of the pyramid starting with identifying the level of sympathy for terrorists among a population. The methodology for understanding the progression to radicalization is contained in social movement theory. While polling data may show that only a small percentage of the population supports violent action, the numbers can equal hundreds or thousands of individuals that support violence within a population. The challenge is in finding and understanding this small minority of the population.

Polling data may also be used to track changes in perceptions and beliefs of target populations. This data is relevant in helping the US determine whether it is winning or losing the war on terrorism. Data to date have shown that the US is not winning the war on terror, not even within the US.

The idea that activism is a conveyor belt to radicalism is inaccurate. Often, activists and radicals may have similar goals; however, they may be in competition with one another to determine how those goals are achieved. They also compete to win the support of the population. Some activists do become terrorists, but to assume that all activists become terrorists is incorrect.

Some popular explanations for Jihadi radicalization include socioeconomic marginalization (frustration); religious fanaticism (Wahabi, salafist); social identity marginalization (frustration); and political grievance. Using polling data, researchers can begin to test these hypotheses.

Mechanisms of radicalization at the individual level may include personal grievance; group grievance (sudden jihad syndrome); self persuasion in action (slippery slope); joining a radical group (power of love); fear or escape to group security; or excitement, status, and money. These mechanisms have been identified through social movement theories and through autobiographies of historic terrorists.

Mechanisms of radicalization at the group level may include extremity shift in like-minded groups; extreme cohesion under isolation/threat; or intergroup and intragroup competition.

Mechanisms of radicalization at the mass level may include external threats; hate and killing by category; and martyrdom. The mass level is a bit of a stretch because hard to make projections from individual or group to mass. However in-group/out-group theory approximates the mass level mechanisms. This area needs more work.

Based on his research, Dr. McCauley reached three main conclusions. First, mass radicalization of beliefs, feelings, and maybe actions can be tracked in survey data. Second, political grievance may be a better predictor of sympathy/justification for terrorism than socioeconomic or identity frustration or religious fanaticism. Third, there is no conveyor belt to radicalism/terrorism: many paths combine mechanisms at the individual, group, and mass levels.

Further research should be conducted on the following elements in order to understand radicalization better.

1. "War on Islam" How does political identification with world umma succeed against nationalism?
2. "Conveyor belt" When does activism lead to radicalism and violence?
3. "Top-down vs. bottom-up" Internet videos build grievance? Chatrooms build trust?

TODD HELMUS

National Defense Research Institute, RAND

Mr. Todd Helmus of RAND spoke about social science foundations of terrorism with particular emphasis on the assessment of factors implicated in individual radicalization.

AMY PATE

Research Director, Minorities at Risk Project, University of Maryland, START

Amy Pate spoke about the Minorities at Risk (MAR) Project at the University of Maryland's START program. MAR is a university-based research project that monitors and analyzes the status and conflicts of politically-active communal groups in all countries with a current population of at least 500,000. The project is designed to provide information in a standardized format that aids comparative research and contributes to understanding conflicts involving relevant groups.

Minorities at risk are defined by several criteria. It includes groups that collectively suffer, or benefit from, systematic discriminatory treatment vis-à-vis other groups in a society. Another criterion is that the group is the basis for political mobilization and collective action in defense or promotion of its self-defined interests. Additionally, the group's population must be at least 100,000 or 1 percent of the country's population for the purposes of this project. The MAR project includes groups who are distinct by reason of some combination of history, language, culture, religion from majority and/or politically dominant group in the state. Groups are

categorized as to type: ethnonational, national minority, indigenous, ethnoclass, communal contender, and religious sect. The project also includes disadvantaged majorities and advantaged minorities.

Groups are considered the highest possible level of aggregation. For example, all Hispanics in the United States are coded as one group. This is one of the more controversial decisions of the project, along with the decision – made primarily for reasons of information availability – to only include groups above a certain population threshold.

The project was started in 1986 by Ted Robert Gurr. It has catalogued over 285 groups and is working on updating the 2004-2006 time period. The project currently has funding to add another 80-100 groups to the database.

MAR is designed to track the political status and activities of 285 ethnic minorities around the world. It focuses on state-group and group-group relations and interactions. The effort develops analytic summaries and risk assessments for all groups included in project. It also maintains chronologies of events.

There are four key variables. The first is political restrictions and discrimination, cultural restrictions, and economic discrimination (annual for 1990-2003). The second is protest and rebellion (annual for 1985-2003) and intergroup and intragroup conflict (annual for 1990-2003). The third is repression (annual for 1996-2003). The last is grievances (biennial for 1990-2003).

The Minorities at Risk Organization Behavior (MAROB) project gathers data on organizations that claim to represent ethnic groups. This effort is unique because MAROB gathers information on violent and non-violent organizations. MAROB collects information on organizations based on the MAR dataset. For example, MAR collects data on Kurds in Iraq; MAROB collects data on the Kurdistan Democratic Party, the Patriotic Union of Kurdistan, the Kurdish Revolutionary Hezbollah, among others.

The selection criteria for ethnic organizations rely on several factors. First, the organization makes explicit claims to represent the interests of one or more ethnic groups and/or the organization's members are primarily members of a specific ethnic minority. Second, the organization is political in its goals and activities. Third, the organization is active at a regional and/or national level. Fourth, the organization was not created by a government. Fifth, the organization is active for at least three consecutive years between 1980 and 2006. Sixth, the umbrella organizations are not coded.

There are four primarily variable "suites". The first is organizational characteristics, which includes type of leadership, ideology, and grievances. The second is organization-state relations, which identified whether the state represses the organization or whether the state has reached an agreement with an organization. The third is external support, either from the diaspora, a foreign state, an NGO, or an IGO. The last is organizational behavior, which includes nonviolent behavior (electoral politics and protest); violent behavior (violent rhetoric, targets of violence, location of violence, type of violent repertoire); and involvement in criminal networks (such as the Unholy Alliance Project).

Some potential questions that MAROB is expected to answer include the following. Under what circumstances do ethnically based organizations use and cease to use political violence as a strategy? What influences organizations' choices of targets (e.g. civilian versus state)?

The MAROB database is being updated by region based on a timeline. The Middle East is complete while Afghanistan and Pakistan are nearing completion. Post-Communist countries are schedule for completion by January 2009. A limited collection of Asian organizations started in fall 2007 while Latin America, Western Europe, and North America are expected to start in fall 2008.

Coders are trained on coding guidelines through use of presentation, short activities, and practice cases. Each coder is trained on finding and validating information. The project is careful to seek coders with specific language skills or regional expertise. The editors supervise coders. They check notes for adequate explanation of coding and check the sources that coders cite. Users are encouraged to send new information to the project as well as request recodes of certain variables.

While the project has been underway for two decades, several challenges remain. First, sparsely covered regions/groups/organizations make data collection difficult. Furthermore, not all regions of the world are equally covered by the news media. Closed areas, such as Burma, make credible data collection challenging. Many areas have conflicting claims of responsibility. Lastly, the coders must deal with ambiguous information.

QUESTION AND ANSWER SESSION

Hugh Mcelrath said there were two things he wanted to hear more about. First, in the discussion of radicalism motivations, the panel did not address the role of fundamentalist religion, which is a thorough critique of modernity. The motivation for terrorism is not about being deprived of socio-economic resources; it is about rejecting all aspects of modern life. Second, it is unclear whether the role of operational effectiveness or competence (such as Saddam Hussein and Timothy McVeigh, which both had operational/military backgrounds) makes individuals pre-disposed to these kinds of activities. The question is that if people have the background and training in operations, are they more likely to use this capability compared to white collar workers? Perhaps datasets like MAROB could look into organizational effectiveness. For example, why are terrorist groups so successful in Iraq when many elements of state and economic stability are missing?

Angela Gyurko suggested that it is not the rejection of modernity that motivates terrorists, but a more subtle sense that what is going on around the individual just is not right. There is a desire to do something to change their environment.

Angela Gyurko also stated that for some people the only path within a particular social group is terrorism.

Clark McCauley suggested that rejection of modernity is not the primary motivation, but the desire to be left alone.

Bob Popp introduced a new topic by asking Amy Pate how her coders are trained and whether automation is a possibility.

Amy Pate spoke about training, automation, and validity. Validity is achieved by trying to find multiple sources that say the same thing. It is essentially a process of triangulation. There are no performance metrics, because there is not a large enough n. In respect to training, coders are trained until they understand and can apply the rules consistently and accurately. Some coders train for longer than others. Some have to be paired with more experienced people. In respect to automation, it has not yet proved possible to automate the coding process due to ambiguities. A computer is not good at determining whether there has been progress or whether an organization has split. What can be done is to use automation that flags areas with low confidence scores for human review. The other problem with automation is that the coders frequently have to look at more than one document to appropriately code a variable. However, automating the collection of sources for review could be useful.

Joshua Sinai returned to the discussion of modernity. He stated that opposition to modernity is one of the causal factors of religiously fundamentalist terrorism and one that he will speak about during his presentation in the next session.

Laurie Fenstermacher addressed a comment to Clark McCauley and Todd Helmus. She suggested that instead of lumping radicals into one pot that it might be useful to look at instigators versus perpetrators. Each group is different in terms of demographics and motivations.

A. Arash said that the Army has access to the Open Source Center as well as databases like Lexus Nexus. The Open Source Center has a huge repository of data that are tagged with metadata and vetted. The Open Source Center faces three problems. First, trying to impose a western academic framework on a target population does not always work. Second, the information space the center is dealing with is full of

publications done by people whose expertise was only developed after 9/11. Third, when experts get too close to a topic, they tend to become myopic.

Brain Meadows stated that the issue of defining a terrorist is important. Eigen was perceived as a terrorist, but led a political career in Britain. Also, if you look at terrorism motivation another way and replace love of country for martyrdom, you come up with many of the same reasons why people join the military to defend their nation. Is there any research that shows any mapping for why soldiers fight and why people become militants?

Larry Kuznar stated that one problem he sees through research on radicalization is the failure to operationalize variables. One might say that socio-economic status is not a predictor, but later contradict oneself and indicate that it is a predictor. MAR is doing a great job of operationalizing its variables.

PANEL 1: GOVERNMENT SPONSORED RESEARCH ON VIOLENT EXTREMISM

Allison Smith, Department of Homeland Security (Moderator)
Matt VanKonyenburg, Office of the Director of National Intelligence
Shawn VanSlyke, Federal Bureau of Investigation
LtCol Lyons, Marine Corps Intelligence Agency

ALLISON SMITH

Department of Homeland Security, Human Factors

Allison Smith spoke about current data collection efforts in radicalization across the individual, group, and community levels. Since this new program area received funding starting in fiscal year 2008, current studies focus on looking at what is already known in the field and how to best move forward.

At the individual level, DHS is laying the groundwork to conduct in-depth case studies of individuals who have engaged in terrorist violence in the US. Case studies in the past focused solely on a few different individuals; this study will include as many individuals as possible. Drawing on social science theories, the effort will begin by examining which variables have already been coded by others and which variables still need to be coded. There will be a workshop this year to talk with people who have already collected data using open source methodologies. In addition, DHS will sponsor research that uses the methods of forensic psychology to understand bombing attacks. Working in coordination with START, the study will look at a cross section of bombing attacks and examine variables related to organizational planning styles and motivations.

At the group level, DHS is studying terrorist group rhetoric with a focus on groups that have engaged in terrorist violence as well as groups that would be considered radical but that have not engaged in terrorism. The study would look at groups with similar ideologies and goals and try to understand why some groups choose violence and others do not. It will examine whether variables coded in rhetoric can be connected to significant attacks by terrorist groups. The work will employ both manual and automated coding systems.

DHS will also sponsor research at the START Center to examine characteristics of groups who participate in IED attacks using such databases as MAROB and the Global Terrorism Database (GTB). IED attacks in the homeland is an area of major concern at DHS. A lot of work in the next few years will be focused on IEDs as a particular tactic of terrorists in the US.

At the community level, DHS is involved in two primary activities. The first is evaluating whether survey data that has already been collected is related to radical activity in the US. The evaluation will seek to understand whether there are any variables in surveys that could be possible predictors of radical activity. The next project not really about radicalization, but is important to homeland security. It is an ethnographic study looking at Muslim communities in US and their perceptions post 9/11 in terms of government activities and integration in local communities. Finally, DHS is working with START on studying counter radicalization programs conducted in five countries including Yemen, Indonesia, Northern Ireland, Columbia, and Saudi Arabia. It seeks to understand the effectiveness of this programs. It will also evaluate whether these programs would be applicable in the US.

All of the efforts listed above will feed into modeling efforts.

SHAWN VANSLYKE

Federal Bureau of Investigation, Quantico

Shawn VanSlyke, FBI, spoke primarily about how the FBI operational unit conducts research.

MATT VANKONYENBURG

Office of the Director of National Intelligence (ODNI)

Matt Vankonyenburg, ODNI, spoke about the Defense Intelligence Socio-Cultural Dynamics Working Group (SCDWG).

LTCOL LYONS

Marine Corps Intelligence Activity (MCIA)

LtCol Lyons, MCIA, spoke about the challenges of operating in the human dimension: Shaping the operational environment and building capacity within the Marine Corps. MCIA understands that the socio-cultural problem is not just about understanding the adversary; it is about understanding ourselves as well.

Shaping the operational environment is not just about killing bad guys nor is it about resolving all disputes. It is about setting the norms and rules so that conflicts can be mediated through non-violent means. If the commander only knows where the adversary is, he will kill him. MCIA is trying to give him other and better options to minimize/mitigate kinetic operations. If war is considered to be politics by other means then politics should also be considered as war by other means. This is particularly important when one considers that many post-conflict groups are now part of the legitimate government.

MCIA has defined culture as the creation, maintenance, and transformation of semi-shared patterns of meaning, sense-making, affiliation, and organization by groups of people. It is a continuing process in which people interact with each other and with their environments. Cultural intelligence is the process and the product of all-source analysis informed by socio-cultural data and by theories, frameworks, and methods from the social and behavioral sciences. MCIA has a range of products that feed cultural intelligence. For example, human terrain analysis helps commanders understand how people interact with their environment (GIS) whereas social network analysis details how groups interact with each other. The purpose of Cultural Intelligence is to predict the likely range of actions for a person or group in order to focus scarce operational and intelligence resources and activities.

MCIA has built a pyramid to help commanders pick the right product for the right mission. The vast majority of Marines need easily accessible and understandable tools. They need to start with broad based framework to give them a cognitive map to help them understand their operational environment. As the product's level of complexity increases, the required socio-cultural knowledge base of the consumer increases as well. The base of the pyramid is foundational data. These are the field culture guides and smart cards that offer a low level of granularity, are available off the shelf, are easily accessible, and easy to use. The middle layer is focused on broad-based mission support. This level includes support for information operations, deep culture studies, and local atmospherics. These products involve a moderate level of granularity, some training requirements, and are more likely to be classified. The top level of the pyramid is tailored mission support, which includes human terrain analysis and social network analysis. These products involve a high level of granularity, are designed for experienced consumers, mission specific and classified.

The method for acquiring information to produce Cultural intelligence is through the concept of the data broker. There is no single data repository for all types of cultural information. Instead of focusing on one super database, the MCIA is determined to make partnerships with a broad range of people across disciplines. It wants to engage with partners in data-sharing, knowledge creation, and knowledge dissemination. The input of cultural intelligence includes human terrain analysis; social network analysis; values, beliefs and narratives; and social organization and culture economy components. The output is that better intelligence preparation equals better decision making.

The next generation of cultural intelligence includes new products. The first is visualizing cultural intelligence by "putting culture on the map." The second is support to influence operations and irregular warfare by

building partner capacities focused on the values and themes of sub-national groups. The third is cultural vignettes, which are short and focused products designed to create a specific cognitive map for each mission set. The fourth is building an analytic capacity by supporting uniquely trained cultural analysts, ensuring 24/7 reachback for deployed units, and creating deployable analytic teams for contingency operations.

QUESTION AND ANSWER SESSION:

Nick Poltorak emphasized the importance of cultural intelligence because diminishing the underlying conditions are very important on the strategic level. It is impossible to diminish these problems without socio-cultural tools and analysis. One cannot make strategy at the top level unless you know what is going on with the human population.

Joshua Sinai spoke about the new congressional initiative on countering radicalization. It is a 24 million dollar effort to establish a national commission and create a new center for the study of radicalization. Some believe it is unnecessary because it is duplicative since other programs in the government and academia, whether in the U.S. or abroad, are already studying radicalization.

Allison Smith stated that DHS was following the initiative closely. She would never say there should not be more research on radicalization; however, there is some concern that there is already one center on radicalization (START) and that a second center might lead to duplication. Further. until we wade through privacy issues related to conducting research on radicalization, we should not spend millions creating a new center. Privacy is really one of the largest issues the community is facing now. It would be helpful if Congress would weigh in on that situation.

Brian Meadows asked whether any data shows radical Muslims in Thailand or elsewhere have similar agendas to their counterparts elsewhere.

Allison Smith answered that she supports the need to look at the rhetoric of what groups are saying. DHS is sponsoring a project to conduct content analysis on two terrorists groups: Central al Qaeda and al Qaeda in Saudi Arabia. Previous research showed that the two groups are similar in values but focus on different adversaries.

A. Arash questioned LtCol Lyons about his use of the word predict.

LtCol Lyons responded that predict in context means anticipating a range of actions – not what is going to happen tomorrow. This kind of prediction helps focus assets.

Kelcy Allwein stated that prediction is more than red teaming. It is really the aspect of looking at our own culture and how can we change from product centered mode into "what if" analysis. It is a very different paradigm.

Patti Morrissey stated that the military has ignored the predictability that can come out of real social science. Last summer when advocating for funds for HTTs, several people with highly technical background asked the team to predict how many people will die on x date. There is a misunderstanding of what prediction means in a social science context. Boundaries must be built.

Elisa Bienenstock stated that when the social science and the defense/intelligence community come together, they come with different perspectives. For example, in social science, in order to understand a dependent variable, you cannot just look at that variable to understand it completely. For example, to understand terrorists, you must understand the climate and environment they come from as well as the people who choose not to become terrorists from the same environment. You cannot just focus on the bad guys to understand them better. Social scientists must be able to move in and out of paradigms and methodological approaches in the pursuit of understanding.

Allison Smith stated that this was an issue that DHS trys to keep in mind. In the rhetoric project particularly, it is looking both at groups that engage in violence as well as those that do not.

Clark McCauley stated that when thinking about prediction, you must also think about visualization. If you think of where you are now as a point, there are many possible routes to the future.

Brian Meadows, who is a physicist by training, stated that there is a way to meet these predictions. The DoD is moving away from determinism. While it is easier to predict where a submarine might be than how a population might react, these predictive tools and capabilities should get more accurate over time. The focus should remain on managing uncertainty in prediction.

LtCol Lyons stated that anticipating the range of actions may mean a commander needs to exercise tactical patience rather than seeking the decisive battle initially because of the uncertainty about what will be decisive. Commanders can make small initial inputs to influence the operational environment in order to gather intelligence on the effects of a particular action and make incremental changes based on the feedback. It is the job of intelligence to highlight the element of uncertainty as much as to reduce uncertainty to support the commander's decision process. In complex environments, commanders should often take small actions, identify reactions and then develop the situation.

SESSION 2: UNDERSTANDING SOCIO-CULTURAL DYNAMICS – TRADECRAFT AND OBSERVATIONS FROM A DISTANCE

Kelcy Allwein, Defense Intelligence Agency (moderator)
Matt Page, O Office of the Director of National Intelligence
Joshua Sinai, Analysis Corporation
Bill Diggins, Gallup
Rebecca Goolsby, Office of Naval Research

Kelcy Allwein, the moderator for session two, stated that in coordination with the Directorate of Analysis at the DIA that the session today consists of a group of experts of modeling human behavior. The session consists of two subpanels. The intelligence community (IC) has taken tremendous strides in creating national open source enterprise. The DIA is still in the process of building tradecraft for open source. Elliot Jardines, the Assistant Deputy Director of National Intelligence for Open Source, said that he wanted to make open source the first source of resort, but not the only source. Open source data provides tremendous information, but cannot do it without marrying all sources together. Sources tip each other off. The speakers today were not chosen because they are terrorism experts; they were chosen because they know how individuals impact the system.

MATT PAGE

Geographic and Civilian Marine, on Rotation as Deputy National Intelligence Officer for Africa, ODNI

Matt Page has an extensive background in Nigeria affairs. He spent the last five years conducting a detailed ethnographic map of the Niger Delta.

The USMC has a long history of conducting operations in challenging environments. Marines were the first troops to be sent overseas. They ventured into Africa many times (Tripoli, Barbary pirates, peacekeepers in Liberia, illegal fishing, etc.). Marines are also military attaches in various countries.

Studies to understand a country and its people must be planned and carried out with nuance. Marines rely on maps to understand these complexities.

Compiling the ethnographic map of Nigeria was driven by the complexity of Nigeria's conflicts, human drama, and exotic landscape. Over last five years, Matt Page had to watch the conflict from a distance as it expanded in complexity and intensity. Page became concerned that conflict was being defined in terms of

Nigeria's energy market. Tribal conflict was simplified as the "curse of oil wealth." The study attempted to develop a more nuanced understanding of Nigeria's protagonists and its people. The study had to reconcile ethno-geographic information and piece together picture of human landscape.

Ethnicity has long been a political question. It affects resource allocation, boundaries, and political formulae. A controversial picture emerged in colonial period. Ethnicity and parochialism was a source of conflict in Nigeria. It is this constant back and forth that keeps the nation dysfunction but also in balance. Any ethnographic analysis will be politically explosive and motives will be questioned, but without objective analysis, Nigeria will continue to entrench ethnic division. Political manipulation of the 2006 census reflects the government's desire to maintain ethnic status quo rather than produce accurate assessment. The census data will be used as political ammunition by some and derided by others. Building credible demographic baseline must be done before lasting geopolitical stability can be realized.

This kind of ethnographic study needs to be done by the Nigerians for themselves. However, the US government can conduct a similar assessment for itself. Analyzing the human landscape is difficult because it is constantly evolving. Maps from colonial times are still being used today as the basis of ethnographic studies. Ethnographic mapping is important in Africa, but mental maps are important as well – especially for Africans. Maps are seen as evidence of British ethnocentrism – putting labels where none exist. Roots of the Nigerian civil war dated back to roles played by ethnic groups during colonial times. Conflicts were caused by groups which were given preference to schools by the British government. A cross disciplinary approach is needed to conduct a true ethnography.

No map is perfect. Accurate maps are based on high fidelity, well researched data. Poor maps are based on anecdotal or incomplete information. They cover up some deeper complexities. A good map must use reliable baseline information because inadequate data will come back to haunt you. This is done by authenticating names of areas by using a wide range of indigenously produced maps, oil surveys, and other sources. Once the data is collected, one can start constructing ethnographic layers. Matt Page conducted field work in 2005. He reached out to Nigerian colleagues, newspapers and magazines which describe people and places (if not events) accurately. The work of anthropologists and cultural historians is also useful as were 20th century ethnographic surveys. Indigenously produced ethnographic sources are useful, but potentially biased source. They are primarily useful for identifying tribes and borders. Additionally, there are some websites describing ethnic group and areas.

Ethnicity is extremely important in the delta. Using maps helps researchers understand the cultural geography of conflict. Chorography is the study of distribution of populations living in area and causal relationships between them. Maps can help explain conflict, identity, and interaction with oil companies. Maps can lead to chorological analysis of demographics, migration, language, trade, economic development, trade and fishing, age groups, impact of central authority, connections between groups, party allegiance, and political networks among others.

Ongoing chorological analysis continues to reveal causal linkages and patterns of political and social behavior that researchers were previously unaware of. The study resulted in a more comprehensive and nuanced view of Niger delta. This information is really welcomed by the community. Mapping the human terrain is needed to understand these complex problems. Without maps, it is difficult to put events into context. This is prerequisite for US success in Africa.

Matt Page stated that the US Human Terrain System (HTS) efforts today have been somewhat disappointing. The effort has failed to create a focus of effort on human terrain geography. Additionally, there are too many people working in an area where no one has shown real capability. The government needs to reflect honestly on progress made to date. Funds are squandered on efforts to build technological solutions, where what is needed now is an effort to create fully integrated, interagency effort to design area expertise, and innovative GIS techniques.

The intelligence community should pull its resources together to create human terrain institutes staffed by analysts, professors, and intelligence analysts. It could work under the Open Source Center. Its objective

would be to understand clients, neutrality, and close links with map services center. Other agencies that want to participate can do so by staffing the agency with billets or to give it funds. With so many disparate efforts, it is time to focus as a community.

JOSHUA SINAI

The Analysis Corporation

Joshua Sinai, The Analysis Corporation, spoke about identifying socio-cultural drivers in religious fundamentalist terrorism. He asks the questions, what socio-cultural drivers make fundamentalists resort to terrorism? What are the tipping points from extremism into terrorism?

Socio-cultural drivers propel different types of terrorists. The drivers are shaped by different socio-cultural dynamics and processes that drive an individual from non-violent extremism into terrorism. Radicalization is the process of becoming an extremist. The radicalization process can occur in either a top down or bottom up approach, as discussed during the Clark McCauley presentation. Different socio-cultural and psychological factors may propel people to become terrorist "managers" – those that have no intention of giving up their lives – or members of a combat team – individuals who are willing to give up their lives for a cause.

The first driver is posited by Jerrold Post, a political psychologist at George Washington University. He writes about generational paths to terrorism. Loyalty to one's parents and the parents' relationship to the regime is a determinant of whether an individual becomes a nationalist-separatist or a social revolutionary. However, the problem with this driver is that it does not account for today's most lethal terrorists: religious fundamentalism. Religious fundamentalists are the opposite of social revolutionaries because they want to maintain a traditional way of life, not change it. Regardless of whether one agrees or disagrees about the role of parental relationships in producing terrorism because other factors play more crucial roles, the third category of religious fundamentalists needs to be considered.

Today's terrorists are driven by a radicalization process that has little to do with parental interaction. Individuals are radicalized by society through media, schools, or other groups. In other cases, such as al Qaeda, the process is bottom up and indirect. Recruitment often occurs through friends, family, mosques, and self radicalization over the Internet. A more relevant model of drivers is Marc Sageman's four prong process[1] on how individuals are mobilized as soldiers for Islam.

In Dr. Sageman's model, terrorists view themselves as heroes fighting for justice to transform societies. Recruitment is facilitated by a four prong process. The first prong is moral outrage over the suffering of Muslims. The second is the interpretation of outrage in context of larger world. The third is resonance with personal experience. The fourth prong is mobilization by networks which takes the recruit to the next level. Joshua Sinai proposed a fifth prong between one and two: the influence of radical Islamic text. While this model is relevant at the micro level, analysts must also look at macro level socio-cultural drivers.

The al Qaeda brand of religious extremism advocates the imposition of Sharia law over society. This is not limited to al Qaeda; Hamas is also implementing Shari'a strictures over Palestinian society in the Gaza Strip. Islamist movements in Western Europe also seek to impose Shari'a laws over their communities. These types of religious fundamentalism represent a rebellion by people who come from traditional societies against modernization. Their rebellion is also fueled by problems in socio-economic status, lack of political rights, problems in integrating immigrants in Western societies, and conflicts around the world. The Internet and mass media provide a means for individuals in one region to be traumatized by conflicts in other regions.

[1] Marc Sageman, "Leaderless Jihad: Terrorism Networks in the 21st Century." University of Pennsylvania Press (December 2007).

Traditionalism is represented by a socio-cultural religious component that promotes gender inequality, creationism, emphasis on religiously based scientific inquiry and knowledge and other traditional beliefs and practices. Traditionalism also promotes a belief in a blissful afterlife, which facilitates the practice of suicide martyrdom . Modernity promotes change in all aspects of life, including secularism, democracy, civil laws and human rights, plurality of beliefs, right to dissent, universal suffrage, social mobility, gender equality, scientific inquiry, innovation, technological and economic development – these are all characteristics of a modern society.

The socio-cultural causes of religiously fundamentalist terrorism (which is not only Islamic but affects Jewish extremism, as well) are, in such a way, driven by the conflict between traditionalism and modernity. The small minorities that engage in such extremism and terrorism refuse to come to terms with the modern world because of the hold of traditionalism over their daily lives and beliefs. These extremists, inspired by ideologues such as Sayyid Qutb, view civil society as incompatible with Orthodox religion. Many countries co-exist with modernity and no country is completely modern. The ultra traditionalists, however, refuse to accommodate themselves to the modern world.

For insurgent leaders, such as Usama bin Laden, terrorism is perceived as the best way to achieve their objectives, since conventional tactics have no chance of success against their more powerful adversaries. For this, they can usually draw on large number of followers.

Opposition to modernity is not the sole reason for religiously fundamentalist terrorism. But it is one of the major socio-cultural drivers. Research needs to focus on micro-level drivers to serve as preconditions for the final shift from radicalization into violent extremism. There is a need to focus radicalization study on the end stage: the final leap from rational behavior to violent extremist activities.

BILL DIGGINS

Gallup Consulting

Bill Diggins, Gallup Consulting, spoke about radicalism in the Muslim world: predictive factors and triggers.

REBECCA GOOLSBY

Office of Naval Research (ONR)

Rebecca Goolsby, ONR, is a cultural anthropologist with a deep knowledge of ethnographic field training. She spoke about studying unsettled times: socio-cultural research from a distance.

The DoD would like to find gold standards for indicators of phase-shifts in political stability of a nation or a region. The primary problem is that these indicators vary from society to society. The second but not inconsiderable problem is that technology is changing the way people communicate, speeding up conflict and flash points. The question remains: How do we forecast conflict, violence, and expressions of public outrage? How do we do that in places we cannot go?

During World WII and after, many anthropologists and social scientists were funded by the Department of Defense to study the cultures and societies of adversaries, especially Japan. Margaret Mead and Ruth Benedict were among the hundreds of scientists who received Department of Defense funding.

Culture at a distance methodology from 1942 mostly comprised of several sources. Researchers studied travelers' accounts, ethnographies, histories and other second hand sources including social science works. They conducted interviews of nationals in the home country. They also studied literature, music, art and ritual, as well as religious works. Content analysis finds its roots in these attempts. Benedict's work The Chrysanthemum and the Sword, based on her study of the Japanese, is still in print.

Post World War II, the DoD funded a great deal of social science research—only a small part of it was "culture at a distance". They focused instead on models such as game theory. Their main interest was in political instability: its causes and indicators. The Camelot Project was brought down largely because of the appearance that the DoD would use this knowledge to meddle in international politics outside of its mission. This led to the creation of the NSF.

There is a problematic nature of new missions. DoD Directive 3000.05 directs that socio-cultural data collection will be everyone's problem. It provides for the expansion of the services' mission from warfighting to managing security, stabilization, transition and reconstruction (SSTR) operations—a huge area in which they are to be the lead, rather than helpers to USAID and others. This builds on the development of the concept of the "three-block war". The DoD does not have the knowledge and understanding of non-Western cultures and societies necessary to execute these missions; hence new research funding initiatives.

Because of this, there are several new culture at a distance methods. Computational modeling has become the sine qua non of culture at a distance methods, but it suffers from considerable handicaps. Its emphasis on computational engineering leads to "everything looks like a nail" thinking. The problem with models is that many do not scale gracefully from explaining small scale issues to larger scale ones. Models are culturally and socially ignorant, which is frequently the case when engineers develop "new cultural theory" with little or no training in the fields of social science.

However, the key problem with computation based modeling is the data. In the development of computational models, modelers continually do not have data that is:

1. the "right size" for their model
2. the "right kind" for their model
3. of the right "freshness date" (mixing 19th century social structure with 21st century historical data is simply wrong)
4. of a sufficient size, richness and robustness for the model

These data limitations bring us back to 1942 methods. Other methods are available. One method involves using case studies. One example involves splinter groups and military arms, which are historically very important in the emergence of violence. See "Modeling Islamist Factionalism" by Michael Gabbay (with consultancy of Mohammed M. Hafez, author of "Why Muslims Rebel"). More work on existing political, religious, and social groups, including historical work, would be valuable. Agent based models may be valuable when based on case studies written by various scholars, who can then also be interviewed and collaboratively engaged. Case studies can provide insight into the traditional places, methods, and reasons that a society has violently destabilized (or peacefully changed, which may need some study as well). However, the dangers of relying on case studies include that the research is often dated and of various vintages (need to account for global factors). It is hard to get data from different societies to fit into the same computational model. Lastly, theoretical bases of cases may be hard to compare.

New methodologies also rely on the internet as a source of information. Studies of the people involved in groups that turn violent show that there are many kinds of people across a spectrum of social backgrounds. One fear is that the Internet is improving the tradecraft, connectivity and incentives to "go local" with violent acts. More creative work to study the Internet and emergent groups is needed. Language and discourse research is greatly needed.

Additionally, the ease and velocity of international travel is allowing researchers access to refugees and sources of information like never before. Today, people move great distances, providing new people to interview (refugees, exiles). New immigrant communities grow up quickly. Researchers can spend days, weeks or months doing short research excursions; multiple trips in a single year are even possible.

Researchers are also engaging in limited ad hoc fieldwork. Local social conditions and local cultural histories appear to be significant in the creation of "hot spots". Scott Atran and Marc Sageman traveled to home

towns of people involved in a terrorist incident to talk to their neighbors about them. They could only do "pick up" conversations with town folk (no formal research permission given). Soldiers collect a variety of ethnographic data. Ad hoc fieldwork is not traditional fieldwork, in that it is brief, opportunistic, and partial and provides a more superficial picture of society. The problem with ad hoc fieldwork is multifold. It provides a journalistic rather than scientific point of view. Can be very small scale and impressionistic. It provides dangerous settings. It does not provide enough data, of sufficient quality, for modeling though it can augment data from other sources for modeling.

To yield the best results, ad hoc fieldwork needs to be done by professionals who can best assess what data needs to be collected, how to get the best quality, etc. It must be sustained over a long period of time to get the best value. And it must be planned to provide data that will validate, course correct, or inform other distance methods.

To understand complex social problems, from drug warlordism in Afghanistan to mass killings in Sudan, it needs to be better understood from a systemic viewpoint. There are two ONR Multidisciplinary University Research Initiatives supporting this task at George Washington University and Carnegie Mellon University.

The biggest problem, bar none, is getting a theoretically sound conceptual model bolstered by adequate and sufficient data that can be validated to some degree. There are few models of military relevant problems that can make this claim. The objective is the creation of a robust, usable toolset for analysis, planning, and forecasting. Computational tools seem the best hope for this.

The research at ONR supports sound theoretical foundations such as computational models and novel data collection plans. These are supplemented by limited ad hoc field studies where possible, advisable and necessary.

QUESTION AND ANSWER SESSION:

Larry Kuznar asked Rebecca Goolsby whether soldiers can adequately capture socio-cultural data they receive through immersion.

Rebecca Goolsby responded that she is not against the idea, but it does not meet the standards and best practices of academic research. The kind of data gathered by soldiers is good information, but it may not be adequate for computational modeling.

Larry Kuznar stated that soldiers are a lost resource.

Rebecca Goolsby replied that the soldiers remind her of travelers in times past who were very educated, but still wrote about the heathen Chinese. It provides interesting data, but is filtered through personal views. The perspective of the soldier is pragmatic; they have an operational point of view. A scientist is more likely to take a broader view. Soldiers are interested in a particular problem set. They will collect data that seems relevant to them and their problem. A scientist may have an understanding of political violence in a number of countries and how it plays out in different areas. Scientists will use theoretical concepts to figure out how the environment is progressing.

Kelcy Allwein asked Bill Diggins whether he had considered media mapping – looking at what the voices in media mean in terms of state government and radical voices.

Bill Diggins responded that Gallup has looked as some media outlets to see what it find most credible. However, it has not categorized them.

Nick Pultorak said that he found out that sympathies toward radicalization were greatest near where broadcasts could be received than in remote areas. This is an area where people can see the outside world and know that they do not participate in it.

Brian Meadows asked how you guard against mirroring our impression against questions asked in polls and surveys.

Bill Diggins responded that they first attempt the question and then work to make the outcome variable more sophisticated. This information is not yet geospatially represented, but it is under discussion.

A. Arash stated that in reference to the media environment, most radical groups are non-state actors. Therefore, they do not have access to media outlets the way we understand it. Technology has created a situation called democratization of information. Anyone can publish without filter. We need parameters to map or find these things. We need to go from gathers of data to hunters of data.

Hugh Mcelrath stated that objectively, traditionalists are right – they are under mortal threat. We want them to go quietly and retreat into an enclave. In order to defang radical traditionalism, you need to let them win someplace. They won it in Iran and the youth is fed up with it.

Joshua Sinai stated that one of the problems with ultra traditionalists is that they want to impose their orthodoxy over others, as opposed to wanting to be left alone They have ambitions and seek to impose their will over others. This is the problem facing Israel and Palestine – in both entities secularists are being threatened by religiously fundamentalist traditionalists.

Bill Diggins stated that when you survey across a broad swath of people, you can find demographic breakdowns in the data.

Bob Popp asked Matt Page to better explain his negative position on HTS given that HTS is less than a year old and has deployed six teams, which is an extraordinary achievement. Bob stated that Matt's proposal is to do what is done in DC all the time – to create another bureaucracy. The reality is that it hardly ever works.

Matt Page qualified his statement by saying that the government needs a program that is far more Darwinistic and that will absorb programs that are successful. However, some programs, especially technologically based ones, do not provide added value. He recommends making these efforts a meritocracy rather than throwing in the kitchen sink.

Kelcy Allwein stated that HTTs are extremely tactical, and they also have a mission to do strategic assessments in many other countries than Iraq and Afghanistan. What we need is to consolidate at strategic level to provide appropriate resources. What the government does not do well is pulling together. Agencies frequently stove pipe data. We need to pull data together to understand global trends.

Patti Morrissey stated that as a USDI contractor at the center of the bureaucracy issue that she thought Matt's proposal was excellent. The HTTs are providing a band aid function today. The US went to war with an Army that did not know how to do SSTR. Some organization needs to bring together GIS, open source, and social science to make sure everyone knows what is important and what everyone else is doing.

Elisa Bienenstock stated that one of things that strikes her is that we have social scientists talking and trying to move forward. There is remorse that we did not realize the importance of social science at the beginning, but does not think that building a bureaucracy will solve the problem. What is needed are standards and ways to communicate. One problem is that the people in charge of socio-cultural programs do not see the big picture and have no experience in social science. You need to have knowledgeable people in charge to ensure than the bureaucracy is not carrying over legacy projects that did not work. If we continue to move ahead without change, we will propagate past problems.

Jennifer O'Connor stated that she believes that the DoD must really reach out to the Muslim communities in a positive way. The DoD can use survey data, such as from Gallup, to understand what works and what does not.

Sue Numrich stated that the problem with getting everyone on the same page is that if it is not the wrong page today, it will be the wrong page tomorrow or the next day.

Joshua Sinai stated that outreach initiatives are very important, and the most effective might be "out of the box" initiatives. For example, Muslim comedy groups present a constructive, non-stereotypical image of American Muslims. What is needed is for someone to write a book about successful Muslims in America or to research Muslim-American Medal of Honor winners, or highlight Muslim heroes.

PANEL 2.1 SOCIO-CULTURAL TRADECRAFT AND OPEN SOURCE COLLECTION REQUIREMENTS MANAGEMENT

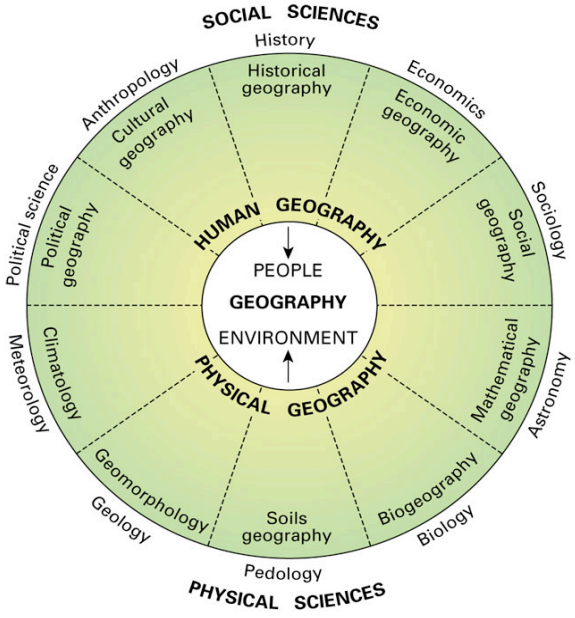
Lizbeth Sydnor, Defense Intelligence Agency
Scott Swanson, Delphi International

LIZBETH SYDNOR

Cultural Geographer, Joint Warfare Support – 2, Operational Environmental Analysis

Lizbeth Sydnor, DIA, spoke about the human dimension of socio-cultural tradecraft. Military operations are inherently geographic in nature. Any credible response to a variety of missions requires an understanding of the physical and cultural geography (human terrain) of the environment since prevailing conditions and local populations can enhance or constrain operations. As the military geography branch of DIA, it is the mission of the Joint Warfare Support Operational Environment Analysis unit is to initiate, direct, and coordinate all source geographic and urban analysis and its effects on the full spectrum of military operations in peace, crisis, and war. Its customers include Office of Secretary of Defense (Policy - Strategic), Combatant Commanders (Operational - Planning), and Warfighters (Tactical).

Cultural geography is an integral element of socio-cultural dynamics as it studies cultural traits, spatial variations among cultural groups, and the human organization of space. It focuses on describing and analyzing the way cultural phenomena varies or remains constant from one place to another. Cultural geography allows us to map the human terrain and understand the factors that shape and influence an operational environment, thus understanding the spatial variations among cultural groups and the spatial functioning of society. Urban geography is the study of how cities function, their internal systems and structure, and the external influences on them and the study of the variation among cities and their internal and external relationships. The human geography areas of interest include economic, demographics, politics, culture, combined social implications, and technology/infrastructure.



The wheel on the left describes the multidisciplinary components that make up cultural geography. The scale of analysis reaches across a broad section: global, supranational, regional, state/national, and local (provincial, district, urban areas, and rural settlements).

The primary areas of focus include:

- Cultural phenomena that may vary or remain constant from place to place
- How humans function spatially
- Geo-spatial distribution of cultural groups, sub-cultural groups,
- Identifying lines of contention
- Identifying contested spaces
- Spatial inequalities and variations among population groups
- Urban cultural terrain

- Economic and political geography and relation to local populations
- Population densities, demographics

Researchers take a multidisciplinary approach and use sources from a variety of fields including:

- Cultural Geography
- National Identity Studies
- Cultural Anthropology
- Religious Studies
- Linguistics
- Narratives and Travelogues
- Ethnic and Minority Studies
- Subaltern Studies
- Economics
- Sociology
- Psychology and Behavioral Studies
- Demography
- Public Attitude Surveys
- Subject Matter Experts
- Raw or Finished Intelligence
- Academia
- NGOs
- Census Data
- Commercial data (market research)
- Media (foreign and domestic)
- Advocacy Groups
- Government Agencies
- Archives
- Literature
- Blog Sites

Cultural geographers face several challenges. First, experiences in Afghanistan and Iraq have shown that neighborhood level data is necessary for operational and tactical planning (Intelligence Preparation of Battlefield).and collection efforts are not yet established to accommodate this new requirement. Second, there is a lack of reliable data sets. For example, there has been no census in Lebanon since 1932. Third, closed countries, such as North Korea, pose a problem. Fourth, no city level data exists or it is difficult to obtain. States are reluctant to release this information.. Fifth, there is a lack of current academic studies in specific countries of interest. This is due to a variety of reasons: security risks inherent in fieldwork in countries of high interest, and Many of the countries of interest pose security risks, and many graduate student

However, there are solutions to some of the problems. First and foremost, there should be appropriate funding for data acquisition. Someone in the community, perhaps the Open Source Center, needs to be designated as principle collector of demographic and cultural related datasets from a variety of sources such as: census departments, market research firms, academic institutions, NGOs. Second, there must be an improved sharing capacity across domains and organizations.

SCOTT SWANSON

Delphi International Research

Scott Swanson, of Delphi International Research, stated that that rigorous academic problem solving requires a holistic approach; however, the military needs to take a bad guy approach. Military units typically have less than two months to conduct a study of a particular problem area. Delphi can provide these teams with 80 percent of the material they need in either a strategic or tactical approach.

Emerging violent non-state actors use or share a common story about their involvement to motivate and empower their groups, and attract and mobilize their audiences. The degree by which infringements are interpreted, personified, and perceived is the variable to which violence emerges. It is typically a conflict between the haves and the have not's over a particular issue. The haves or have not's can be anything. We often see the haves as being the minority population with the have not's being a majority "minority" population. The conflict becomes even more heated when the have minority are seen as ethnic minority.

Within the Process Model of Socio-Cultural Dynamics, Group Disposition Analysis brings Operational Net Assessments to Intelligence Net Assessments for a significant piece of the Common Operational Picture (COP). Joint intelligence, surveillance, and reconnaissance system supports situational awareness.

Psychological Behavior Assessment consists of a proprietary Belief Model and use of Social Cognitive Theory to assess behavior based on an individual and/or group's belief system, notion of reality, and environmental interaction. Strategic requirement is the problem definition or "intelligence synchronization" which was formerly known as "collection management".

The model relies heavily on initial requirements. It is critical to identify what users actually need, their mission, and their goals. It requires a four prong approach: anthropological research; open source collection; human intelligence collection; and environmental research. The model will synthesize that information to assess behavior.

Complex adaptive systems include nature, society, and science. Socio-cultural dynamics, which involve tradecraft and observation from a distance, include society (ethnographics/irregular warfare intelligence); individuals/groups; relationships/interactions/associations/means of organization/traditional methods of mobilization; and external stimuli (laws of cause and effect).

Collection work must consider the high and low levels of influence to discover all that is between. The human terrain has been shaped by the entire system including the political, infrastructure, information, social, economic, and armed elements spheres. You must understand on what kind of person you are dealing with and how they fit on the shaper/enabler/sustainer pyramid.

The shapers of human terrain include the critical yet less examined elements of intent (mind) and will (heart). IW intelligence utilizes complex analysis to discover these elusive battle-space elements to target key nodes and links. Its history is in anthropology, the basic psychology is individual and group and the current behavior is intelligence. Operational considerations differ by the level of strategic and tactical relevancy and context. For example, where is the line between the have's and the have-not's? How can they shape the transformation?

Enablers and sustainers are involved in collection and analysis. This includes logistics, tools, time, demand, supply, processes, networks/partners, environment, ethnicity/language, and historic reciprocal obligations. Many of these nodes and links are difficult to penetrate due to the barrier of entry. The lockout mechanism is built into the system through ethnicity/language and reciprocal obligations.

Once mapped, relationships can be better identified to target the COGs as opposed to opportunity targets at the lower—less significant levels. Operations are less focused on unpredictable / insignificant 'affiliates', which drain our resources and soldiers' lives.

PANEL 2.2: REMOTE OBSERVATION USING LARGE DATA SETS

Hugh McElrath, Pacific Northwest National Laboratory
Robalyn Stone, Social Science Automation

HUGH MCELRATH

Pacific Northwest National Laboratory (PNNL)

Hugh McElrath of PNNL runs a high security computer forensics lab for the DoD. He is a retired Navy Seal intelligence officer. He tried to get people to use their brains, not computers, to influence people and groups.

The nexus for computer forensics is addressing large data sets. It is important to develop tools to enable researchers to develop these kinds of sets. Bioinformatics, which is computationally powerful, is used to look for malicious code in operating systems. There is also a techno-social group out there. One piece of work in this arena is working with open source information and applying analytic tools to it. Apparently, it is novel to use same tool to address open source and classified datasets and merge them. This program was

demonstrated to DIASPO. INSPIRE was developed at PNNL to allow an analyst to look at a large number of text documents and look for words and phrases. The program graphically portrays the output. A subsequent project is porting INSPIRE over to a high power computing environment.

Another open source tool is project Argus, which is a CDC early warning on emerging diseases. Currently, the system is updated manually by looking for mention of people getting sick in certain way. PNNL wants to provide automation to process. They want to look at larger set of data beyond media reporting including blogs. It would look for evidence of social disruption.

ROBALYN STONE

Social Science Automation, SSA

Robalyn Stone stated that SSA works as a content analyzer for very large datasets. The tool makes coding schemes in the process of creating large data sets form large texts sources. It has an automatic generation process that is transparent. From this data, the analyst produces dashboard like metrics to describe the data. This has created an impetus to create measures of effectiveness and validation.

The SSA team worked on political speeches. The manual collection was undertaken by undergraduates. The team then moved on to different genres: therapy text, Mars Bars, and online media. They used Lexus Nexus because it was formatted in a consistent way. The tool quantified how many words in the texts were indicative of mistrust.

The team discovered that people and computer coders make different types of errors. Humans make boredom errors. Their accuracy always goes down over time. However, if a computer sees a new word, it will not code it. The operation code measures how you present self and others in text on cooperative or conflictive dimensions. To do this kind of analysis, you must have a strong methodological process in order to get strong data out.

During the Integrated Crisis Early Warning System (ICEWS) effort, SSA used BBC Monitoring and downloaded documents relevant to ICEWS. They created lots of data to track what happens in newsfeed and compared it across countries and across time to determine the overall newsfeed on a country in terms of its violence or stability.

Using an example of a three month period in Iraq immediately after the invasion, there were three primary groups in the country. The tools compared how distrustful their language was and evaluated their view of the political universe. They found that the Sunni's had the most distrustful rhetoric.

SSA also conducted a similar study with terrorists and nonterrorists. The more negative he terrorists, the more positive the nonterrorists are. They are starting to do a similar project in modern standard Arabic using Profiler Plus, which allows you to do searches in the original language and do a selective translation. This kind of analysis can be done very quickly – many times in a matter of hours.

The tool can also extract and compare entities for distrustfulness. They use a combination of tools to look for themes and people. It is a pragmatic way to build structure while reducing the analysts' reading load.

QUESTION AND ANSWER SESSION:

Robalyn Stone responded to a general comment about how distrust words are identified. She responded that there is a coding scheme. However, a decision must be made by a human to determine whether the word is distrustful or not. They train coders carefully.

Brian Meadows asked if SSA can handle blog text.

Robalyn Stone responded that it is possible, but challenging. Blogs often use casual language as well as the informal you. It is also difficult to identify which part of a blog page is actual text of interest and which part is supplemental. However, once you format a particular page, you can work with it.

Kelcy Allwein asked Lizbeth Sydnor whether she can conduct urban terrain analyses from a cultural aspect – such as urban tribes. What do you find that is different between an urban study versus state or non-state actor.

Lizbeth Sydnor replied that it depends on research questions. The urban environment involves finer grain data.

PRESENTATIONS DAY TWO

SESSION 3: UNDERSTANDING THE HUMAN TERRAIN – TRADECRAFT AND STREETCRAFT METHODS

Patti Morrissey, Support to Office of the Secretary of Defense
Kathleen Kiernan, Kiernan Group

Kathleen Kiernan opened session three by talking about tradecraft requirements in the human dimension. She spoke about the spectrum of tradecraft, from Satellites to Shoe Leather. She recommended that the intelligence community put more emphasis on shoe leather part of collection and analysis. Additionally, you do not have to steal information for it to be valuable. Kathleen then introduced the speakers for panel 3.1.

PANEL 3.1: APPLYING LAW ENFORCEMENT METHODS TO GATHERING HUMAN TERRAIN

Gary Greco, Defense Intelligence Agency
Michael McNicholas, Phoenix Group
Jim Powlen, Logos Technology
Rich Shimon, Department of Commerce

GARY GRECO

Defense Intelligence Agency

Gary Greco, DIA, spoke about the Elements of Style of tradecraft. In deconstructing the element of managing analysis, you have to take what you know instinctively after years of analysis and present it again. Ten years of managing a terrorism analysis organization has lent Gary Greco some authority to speak on the issue. He has long been involved in the recruitment and hiring of analysts. This presentation will provide quick, no nonsense advice on how to manage analysts.

An appropriate balance of skills, knowledge, and abilities need to be incorporated into an analysis team. It needs to be half Army Ranger company and half artist colony. High GPAs and study abroad experience is a good, but not sufficient criterion for a good analyst. Good analysts are adaptable, intellectually flexible, and have a high degree of comfort with ambiguity. Often, it is the people who went to school full time, worked as a waitress, and still managed a good GPA who make great analysts. Similarly, former soldiers who work nights to put themselves through college also show great potential.

An analyst's job is simple: they read, write, and present. In the development of a good analyst, no preparation matches the analyst's comfort level with ambiguity. One never possesses all the knowledge necessary to make a call. It is critical for analysts to understand the relationship between all information needed and all information wanted.

Furthermore, the depth and breadth of an analyst's rolodex is more important than personal understanding. It takes a minimum of 20 months to develop all skills necessary to be good analysis. These analysts are not trained to be regional experts, but they must be familiar with the region and understand the customer's needs.

Analysts need to develop skills to have conversions with their professional counterparts. Frequently, analysts get into useless arguments. This is a bigger problem than one might think. The first indicator of this problem is the analyst who does not receive constructive criticism and edits well. Analysts need to set ego aside, discuss differences of opinions, and work well with others.

Analysts should have experience with analytic tools since no one tool can provide everything an analyst needs. Something an analyst learned in the back of a Landrover using primary tools may still work. All that is required for analysis is having some way of writing and saving. Even a Word document can be an effective tool.

Analysts must also exhibit finesse - the practice of combining knowledge, thought, and intellectual stamina. Personal conviction of paying attention to a few small details may have an enormous impact on outcomes of intelligence. Finesse can be that extra element that brings something to edge of excellence. Analysts should have an aptitude for knowledge. They must combine perception with delicacy in regard to emotions and feelings. Presenting briefs with finesse to senior managers can have life and death results. Analysts with finesse understand issues with better depth and knowledge than other. Good analysts are those who care as much about operational environment as producing a final product.

MICHAEL MCNICHOLAS

Managing Director, Phoenix Group

Michael McNicholas, Phoenix Group, spoke about threats to the commercial maritime cargo supply chain OCONUS. Because the company's goal is to protect ships and port, it has had to establish offensive collection and operations.

Approximately 90 percent of the food we eat, clothes we wear, and products we use are imported daily via commercial maritime transport. Threats to ships and ports include drug smuggling, stowaways, and terrorists.

The initial links in the supply chain include the exporter and transit to the seaport, at the seaport and during loading of the ship and transit to CONUS. Drugs are often creatively hidden in cargo, such as drugs stashing inside individual bananas. Drug couriers can also be hidden away in cargo shipments. These could easily be terrorists or terrorist materials. According to a US Coast Guard report, 25 Islamic extremists entered the US onboard commercial cargo vessels. Narco traffickers are more frequently taking consignments of extremists to the US. There is an increasing level of cooperation between MS-13, narco traffickers and Islamists.

Drug traffickers and smugglers quickly adapt to new countermeasures. Shipping container seals can be ingenuously removed and reassembled to be easily removable for future runs. The ship has many areas vulnerable to traffickers, smugglers, and stowaways. These include waterside, anchor chain, mooring lines, rudder, gangway, and inside and on top of containers. Smuggling can also be accomplished via hull attachments.

The greatest challenge to cargo supply chain integrity is at the production phase and until the ship departs from the load seaport. Mitigation efforts include human intelligence (HUMINT) collection at OCONUS locations. HUMINT operations might include penetrating an organization or recruiting a lower level individual to find out interdiction information. The human collection component is critical to stopping these illegal activities because Customs cannot provide the level of analysis needed.

RICK SHIMON

Special Agent in Charge, Washington Field Office, Department of Commerce

JIM POWLEN

Logos Technology

Jim Powlen, Logos Technology, was an Army CID Special Agent and a polygraph Examiner with approximately 15,000 interviews under his belt. He talked about human dimension exploitation.

The Criminal Investigation Task Force (CITF) was created in early 2002 by the DOD to conduct investigations of detainees captured in the Global War on Terrorism. CITF became the investigative arm of the DOD after OSD was given the mission to prosecute the detainees. OSD transferred the responsibility to the Army, where the CITF now fall under Army Criminal Investigation Command (CID). The need for analysts and intelligence professionals was huge, especially for law enforcement officials not used to working with classified materials.

A lesson that has clearly emerged is that Combatant Commders need a law enforcement perspective to effectively defeat insurgencies. Terrorist groups often act like organized crime groups. The intelligence community may not have much experience in this, but law enforcement does. Law Enforcement Officials and Intelligence Analysts are trained to process information differently; the collaboration of these two skill sets routinely discovers new significant information that give Combatant Commanders a more thorough understanding of the battle space.

In Iraq (prosecutions have just started in AF, I don't have any information on how those cases have been adjudicated), the CITF enjoyed an 80 percent conviction rate with 97 percent of convictions exceeding six years. Six years is the life expectancy in an Iraqi prison. Other units operating in area enjoy only a ten percent conviction rate. CITF success has come from a training program developed from the ground up with a high degree of flexibility. CITF identifies good ideas and sends it down range to see how it works. Because it is a joint environment, the CITF is not as constrained as other units when considering new ways of doing business. It is also well funded, which allows for innovation.

The CITF has also built a consolidated data base using I2G. It is a more flexible version of the Investigatives and Information Management System, I2MS, used by the Air Force Office of Special Investigations. It is also a relational database. The database is now web based instead of server based. The troops deployed overseas find it much faster to use and it provides CITF Agents and Analysts stationed around the world with complete up-to-date information in real time.

The CITF is also successful because of the quality of its mid-level management. CITF is organized into divisions, directorates, and units. Division leaders are open-minded and question the status quo. They also hold people to high standards of performance.

CITF was recognized early on for its training program. It is unique in military history. The learning curve for newly assigned people was six months to operate effectively. That timeline was unacceptable and was cut to six weeks by the development and implementation of a comprehensive Entry-level training program. Initially a five-day 50-hour program it has been expanded to two weeks. Training revolves around mission and organization, and then gets extremely heavy into culture. The program draws upon the expertise of subject matter experts (SMEs). If agents and analysts cannot understand information in context, it will get them nowhere. Additionally, the CITF's interview protocol is highly successful at educing information from detainees.

In order to be an effective Polygraph Examiner or interrogator, you must understand culture. The two main components are that you must know a lot about the person you are talking to and the organizations or events he is suspected of being associated with.. Interrogation becomes significantly more complex when working through a linguist. Powlen stuck with the process for over two years before developing his own methodology. He structured his interrogations based on cultural relevancy. He relied on strong cultural norms of host/guest relationship. He made the interviewees his guests, which obligated them culturally to reciprocate his hospitality through cooperation. This proved to be very successful. His team was able to get information out of interviewees that military intelligence could not obtain in five months. The method is now formalized into a protocol.

To conduct interviews effectively, the interrogator must put all biases on the side and just accept the culture for what it is. The interrogator must get down to the bedrock of cultural foundation. What they found with Eastern populations is a strong set of social behaviors and expectations that existed pre-Islam (Bedouin Ethos as termed by Rafael Patai). The interrogators employed a social interaction scenario, which instills

social obligation to the interrogator. When scales are out of balance it is uncomfortable for the interviewee. In a reciprocal society when the scale tips, they want to balance the scale. The only way they can do that is by giving information.

QUESTION AND ANSWER SESSION:

Angela Gyurko asked Michael McNicholas if his teams carry radiation monitors.

Michael McNicholas answered that they do not.

Rich Shimon stated that Commerce prohibits the export of items that the NRC would not. There are disconnects that are a little disturbing. It may or may not be adequate, but it is all we have. He was unsure whether the US has ever married different regulatory regimes before. The penalties are more than adequate, but there are gaps.

Tom Johnson asked Jim Powlen whether his interrogation model has been used for other ethnic or cultural group.

Jim Powlen replied that the answer to that is in his manual. Anyone with a government email address can request it. Hundreds of people will tell you difference between cultures, but no one operationalizes it. The manual breaks it down in great detail. Essentially, we do not distinguish between the social and professional worlds. Everything is social and brings strict social requirements such as using the host and guest model. By being gracious host, the guest is incurring social reciprocation. In a system where influence is power and authority is not, by setting yourself up as an individual of influence, the interviewee wants to put you in his network so you can help him. The interviewees are not afraid of telling; it is a blameless culture.

Tom Johnson replied that it is a culture of revenge; it is not blameless at all.

Jim Powlen stated that in a shame based culture, you have to find mechanism of avoiding blame. There are no repercussions for unknown sins. It is that they will not tell us because they do not like us. We make them like us and break down barrier to information sharing. In back of the manual, there is an appendix about how to develop compliant cooperative populations in detention environments. Additionally, there is not a big population of "true believers." They are actually easy to interrogate because they are proud of what they have done and want to tell you about it. The others are less committed to the cause and the ideology and have concerns about incurring shame from their actions.

In response to a question about how long Commerce waits to pull the plug on an operation, Rich Shimon stated that it is not easy. You can never do harm to national security and nuclear triggers is not an operation Commerce would typically bait. Finding the overarching network does not really work anymore in the internet age. What we find is that 95 percent of cases, whoever is indicted pleads out. They are required to tell all. That is how they flush out the network. If they lie, penalties shoot way up. Commerce has brought in members of the intelligence community to listen to those discussions. They would love to have social scientists to sit in as part of team as well.

PANEL 3.2 APPLYING SOCIAL SCIENCE METHODS TO UNDERSTANDING THE HUMAN TERRAIN

Patti Morrissey, USD(I) Contractor, Human Terrain Lead
Stefan Kaszubowksi, YouGov
Andrea Jackson, Johns Hopkins University

PATTI MORRISSEY

USD(I) Contractor, Human Terrain Lead

Patti Morrissey, USD(I) contractor, spoke about how to improve intelligence support to stability operations, which falls under the irregular warfare roadmap. She emphasized the importance of understanding the human terrain.

The genesis of the human terrain initiative began with policy and doctrine including:

- FM 3-24, Counterinsurgency, Chapter 3: Socio-cultural IPB for COIN
- Irregular Warfare QDR Roadmap, Use social science to benefit tactical operations – USD(I) and USD(P)/SOLIC set up Human Terrain IPT to provide oversight
- DoD 3000.05, Military Support to Security, Stability, Transition and Reconstruction
- GWOT

The initiative was also based on operational experiences including:

- Commanders frustrated by insufficient understanding of the target area population and its impact on operational decisions –
- Tactical commanders submitted Joint Urgent Operational Needs Statements to articulate shortfall
- JIEDDO responded to shortfalls by funding proof-of-concept which provides social science expertise to BCT commanders
- TRADOC managing Human Terrain System proof-of-concept; sending Human Terrain Teams to theater

There are three key elements to the human terrain. The first is developing a socio-cultural dynamics data network and repository. The second is developing data gathering, visualization, and analytic tools. The third is recruiting, training, and deploying experts to support decision-making.

ANDREA JACKSON

Johns Hopkins University

Andrea Jackson, JHU, spoke about social science research and analysis for counterinsurgency. To understand counterinsurgency, you must first understand and define the elements involved. Politics is the competition of interest for power. Insurgency is a local political competition between insurgents on one side and counterinsurgents on the other. They compete for the right to a monopoly on legitimate force – for the right to be the leviathan. Competition is fundamentally about political legitimacy.

The population has an information advantage over outside groups. Those that need the population's acquiescence include any political actor looking for information on who is working with competing political actor or any political actor seeking to ensure population will not turn them in to the competitor. Additionally, others in need of support of the population including any political actor that needs other kinds of support from the population (recruits, food, water, money, safe haven, etc.) or any political actor that wants to rule a territory in the long term.

Political actor effectiveness is the population's acquiescence to authority. Social scientists want to understand this relationship and why populations acquiesce. To do so, social scientists must describe the actions of both political actor and study the effects.

Collecting information for counterinsurgency (COIN) operations involves identifying key groups in society and representing where they are located on a map. For each group, researchers seek to identify security (level, sources, threats); income and services (level, sources, gaps); beliefs and communications systems (narratives, symbols, norms and sanctions); and authority structures and figures (identity, structures, levels of authority).

Social science is the application of consistent, rigorous methods of research and analysis to describe or explain social life. Its purpose is inference: using facts we do know (data, observations) to learn about facts

we do not know (theories, hypotheses). It is used to make descriptive inferences (semantic program); to place or make observations within a conceptual framework to allow people to understand a phenomenon; and to make causal inferences (syntactic / pragmatic program). Because of the large number of potential causal factors for any social phenomenon, establishing causation is difficult. However, if we can establish the cause of a social phenomenon, then we can make policy to affect that phenomenon.

Data collection involves various methodologies such as ethnography: participant observation; surveys: interview processes (formal, informal, group, individual, structured, and unstructured); record or document review; and history: conducting interviews or reviewing documents about the past. Data analysis involves direct interpretation: analysis by an individual's reflection and synthesis; quantitative analysis: using standard methods of statistics to ascertain relationships; and formal modeling: analysis by creating a formal system that mimics the world. The number of cases can be single, small n, or large n.

There is a set of successful methods as defined by Andrew Abbott in Methods of Discovery: Heuristics for the Social Sciences. The first methodology is historical narration in which logical flow of the argument is the focus and which occurs within a qualitative / narrative setting. The second methodology is ethnography and single n narration. This involves participant observation and in-depth interviewing; analysis of data looking for themes and patterns; and hypothesis generation and description. The third is small n comparison. This involves a detailed understanding of same factors in each case; process tracing; and hypothesis generating / hypothesis confirming. The fourth is large n studies / standard causal analysis. This involves standardized measures / coded aspects of large number of cases to produce statistics; generalization to a larger population; uses statistical models to infer something about the relationships between those measurements; numbers allow for comparability; and hypothesis confirmation. The last is formalization / modeling. In this case, no real data is required to create a theory or model; models without data can generate hypotheses; logical flow within quantitative context is the focus; and models developed from empirical evidence to produce models of social behavior for forecasting.

In evaluating which method is most appropriate, it is important to keep in mind that multidisciplinary methods reinforce one another. An example for COIN in Iraq involves identifying levels of authority of actors, describing causes, and identifying points of leverage through:

- Ethnographic field research, which generates description and causal hypotheses
- Interviews and focus groups, which generates description and hypotheses in more generalizable way
- Question design and comprehensive list of likely answers
- Surveys generalize to larger population and geo-locate constructs through geographic sampling
- Surveys provide basis for regression models

Regression models provide predictive planning tools.

Field research in Iraq is particularly complicated. Researchers must use semantic programs and descriptive inference as well as create a pragmatic program with causal inference. Some of the challenges include sampling, reactive bias, and interpretive bias. Coping with these problems is done by working through local researchers, geographic sampling and purposive sampling, redundancy of methods, and redundancy of research networks. On the ground analysis is informed by area studies, topic specialization, and occasional forays into population.

STEFAN KASZUBOWSKI

YouGov

Stefan Kaszubowki, YouGov, spoke about how to stand up a research outfit in a conflict environment. In 2003, he and fellow recruits from medical college and social science departments built a company to meet the information needs of military clients. The company primarily works for the psychological operations community. They conduct qualitative research through focus groups. Quantitative research is ideal but nearly

impossible to stand up in locations such as Iraq and Afghanistan. The studies try to approximate randomness, but they never quite get there.

One of lessons of Fog of War was to empathize with the opponent. It is important in the context of understanding the competing narratives of conflict. This is what the focus groups are hoping to achieve.

Currently, focus groups are being conducted almost daily. It is not generalizable because of small n, but it meets the military's real time requirements. Focus group leadership is primarily accomplished through training nationals. This is especially important in areas where you cannot bring your own nationals. Training frequently takes place in Kazakhstan or Baghdad.

The focus group sessions run on almost an eBay-like model. It is as if each facilitator has a power rating. People come to the focus group because they know the facilitator's reputation. Facilitators take on a certain amount of risk. There is only so much can be done to mitigate that risk especially since there is a limited ability to vet respondents. Each focus group session must have a plan B and a back door. There are often things you cannot account for, such as conducting a focus group under sniper fire. The groups must also be culturally sensitive. Groups are split on the basis of gender. The survey questions must also be culturally relevant.

The researchers are challenged by attempting to approximate random sampling. Many areas cannot be appropriately sampled because they do not have clear boundaries. To try to overcome this obstacle, the researchers use snowball sampling with a random element. The results are usually good for tracking changes in attitude, but it is not good for drawing population estimates.

Patti Morrissey concluded the panel by stating that the point of this was to think through challenges of applying social science techniques in conflict situations. It is an effort in understanding and looking for the underlying theory – the reason that you build models the way you do. The questions are meant to elicit useful data for commanders. When you start putting localized studies together you can put those together to start understanding complex societies and can therefore conduct effective stability operations.

PANEL 3.3: SOLDIERS AS HUMAN TERRAIN SENSORS

LTC Lynda Granfield, DOS
MAJ Craig Gendreau, Army

CRAIG GENDREAU

MAJ 95th Army Civil Affairs Brigade

Craig Gendreau, Army, was in civil affairs for seven years. MAJ Gendreau is an Africa specialist, French speaker, and has four tours in Afghanistan and Iraq. He spoke about tactical collection and fusion of civil information. In particular, he focused on civil affairs support to the BCT during the Operation Iraqi Freedom surge.

Seven years ago, there was only one battalion of active duty civil affairs personnel. They were traditionally used to support Special Operations Forces. They were used in Baghdad for the first time as a surge tool to a conventional unit.

The CA information objectives were to provide multi-dimensional situational understanding, provide situational analysis, and to work on the development of possible solution sets. The focus of CA operations is human terrain identification and prioritization (sphere of influence management) and civil reconnaissance.

Within the human terrain network, there are sphere of influence engagements. These works identified political/tribal leaders throughout the district and their areas of influence to facilitate reconciliation efforts.

They also identified facility managers and public works workers to facilitate transition to Iraqis - fixing Iraqi problems. They also prioritized efforts to optimize limited resources for engagement.

The civil reconnaissance focus includes several important areas. The first are critical infrastructure nodes. These include power plans, water treatment plants, sewer nodes, electrical substations, etc. The second are economic zones including markets, banks, and medium to large businesses. Prior to going in theater, CA do their best to get a baseline picture. Tools have become increasingly important, but they are only as good as the data put into it. Civil Affairs needs a mapping tool like ArcGIS, Analyst Notebook, or something that comes out of the Mapping the Human Terrain Joint Capability Technology Demonstration.

Civil information fusion occurs at the Civil Military Operations Center (CMOC). This Civil Military Operations and Information node refined the concept of operations for the BCT because they were previously not tracking the civil dimension graphically. It also utilized ASK (ANB) to build personality linkages (analysis of relational environment).

In conclusion, there are three critical nodes in the Civil Affair analysis process. The first is collection. Information is the critical base element to the process. Without this, there is no system. The second is consolidation. Analysis and understanding the information makes the process function. It is an art and a science. The third is dissemination - sharing the situational understanding is the key to success. Dissemination must flow up and down in coordination and integration with partners.

LTC LYNDA GRANFIELD

Department of State

LTC Lynda Granfield is on active duty in California. She most recently commanded a Provincial Reconstruction Team (PRT) in Jallalabad, Afghanistan. She spoke about the role of PRTs in assisting the Afghanistan government with counterinsurgency.

Since 2002, when the first Provincial Reconstruction Team (PRT) was established in Gardez City, Paktia province, PRTs have become a key tool for the international community to assist Afghanistan in becoming a secure and self-sustaining republic. They represent, at the local level, the combined will of the international community to help the government and the civil society of Afghanistan become more stable and prosperous.

PRTs, due to their provincial focus and civil-military resources, have wide latitude to accomplish their mission of extending the authority of the government by improving security, supporting good governance, and enabling economic development. This engagement of diplomatic, military, and economic power by nations at the provincial level allied with the wide latitude to accomplish their mission has been a strength, as it provides flexibility of approach and resources to support the provincial government structures and improve security.

The PRT's purpose was to conduct civil-military operations in Nangarhar Province in order to extend the reach and legitimacy of the government of Afghanistan by:

- Promoting good governance and justice
- Enabling an effective Afghan security apparatus through training and mentorship
- Facilitating reconstruction, development, and economic growth

It ultimately created the conditions for self-sufficiency, enduring prosperity, a secure and stable environment in the province.

While the PRT's purpose did not change, the country's understanding of its purpose grew over time. There was a shift from minor goals of small infrastructure and liaison work to finding drivers of instability and working with/through the government to eradicate them. The shift in understanding the mission required a re-alignment of national power to be inclusive as a unity of effort, and a command group or "executive team"

was formed. The shift in understanding also happened up and down the military chain, where USAID and DOS put officers into each RC and headquarter to try and ensure consistency up and down the chain.

In its progress through interagency synergy, the PRT realized that real progress could only be made by engaging all of the key players. These include the PRT, DoS, USAID, Department of Agriculture, the provincial government, Afghan security entities, and the UN Assistance Mission to Afghanistan.

The PRTs used unsophisticated tools to map out the human terrain – primarily because the connection with headquarters and their tools was so arduous.

The PRT's focus and lines of operation fell into three nodes. The first is security, which consists of assessments, training, and mentoring. The second is reconstruction and development, which includes building human capital and improving economic viability. The last is enhancing governance and judicial reform, which including mentorship of provincial and district leaders, outreach to district/village leaders, infrastructure development and improvement, judicial seminars, and provincial coordination centers.

LTC Granville then changed gears and spoke about the Humanitarian Information Unit at the Department of State. The Humanitarian Information Unit (HIU) serves as a USG interagency center to identify, collect, analyze and disseminate unclassified information critical to USG decision makers and partners in preparation for and response to humanitarian emergencies worldwide, and to promote best practices for humanitarian information management.

HIU gets its information from a variety of sources including Relief Web, the United Nations, other donor countries, USAID, and others that provide situational awareness on complex emergencies and humanitarian crisis. Despite high expectation from practitioners in the field, HIU does not have a giant database in the sky. Because the HIU office is so small and the budget is limited it does not produce products normally unless it gets a request to answer a very specific question for DOS and USAID policy makers.

Patti Morrissey stated that the previous presentations demonstrate the wealth of information and ingenuity out there. The objective now is to bring it all together. If you can figure out where it is coming from, show a picture, interpret it, and take action – that is what is needed today.

QUESTION AND ANSWER SESSION:

Hugh Mcelrath stated that it is too bad that what the presenters have done is not considered intelligence.

Patti Morrissey stated that Matt Page's recommendation struck home during this panel. A large problem, which is mostly a leadership issues, is figuring out how to pull everything together. The reason this work is not called intelligence is because the intelligence community has not thought of it as important. That is part of the transformation we are in the midst of.

Andrea Jackson noted that there is a portion of social science research that is not well funded and that is the collection of the data. She also supports Matt Page's comment. There needs to be some kind of organizing structure that makes sure this happens. Right now it is an adhocracy.

Stefan also suggested that Requests for Proposals (RFP) be written more generically. Often the RFP requires a certain technology that may not be appropriate or necessary. A more open system would allow more innovative, tailored responses.

Andrea Jackson stated that it is a catch 22; too loose a RFP begs for poor project work.

WORKING LUNCH: CIVILIANS AS HUMAN TERRAIN SENSORS

Thomas Johnson – Naval Postgraduate School (NPS)

THOMAS JOHNSON

Director, Program of Culture & Conflict Studies, Naval Postgraduate School (NPS)

Thomas Johnson, NPS, spoke about compiling Afghan human terrain data. He has been a student of Afghanistan since 1986. He teaches advanced Afghanistan studies and lectures many of the PRTs. He believes that in order to operate effectively in Afghanistan, one must understand Afghanistan's cultural context and history. An anthropology background alone and in the absence of a proper Afghan context is not always useful. Human Terrain Teams and others need culture specific knowledge to achieve their mission in foreign countries.

Professor Johnson created an open source portal for one stop shopping for socio-cultural information on Afghanistan. It is located at www.nps.edu/programs/ccs. It was deployed six months ago to interact with PRTs – even those with bandwidth problems.

The site collects data of all kinds including genealogical information. It has received corrections even by the people listed in the genealogy.

The portal facilitates a subject matter expert networks and promotes interaction among the key players including DOS, NGOs, USAID, BCTs, and warfighters.

The site drills down to the provincial level (and in some instances the district level); 21 of 34 provinces have been completed. The site provides information on detailed maps, refugees, education, health, topographic information, etc.

Tribal genealogy is an extremely important dynamic addressed by the site. There are over 40 tribal trees of ethnic, tribe, and clan divisions.

The site has current and easily accessible publications on culture and operations in Afghanistan. It produces a bi-monthly, peer-reviewed web journal. The site can also received requests for research to facilitate HTS reachback.

In Afghanistan, drugs account for 50 percent of the economy. The police are 90 percent corrupt. The state has the lowest tax revenues in the world and it ranks 195 out of 197 in poverty rankings. For all of these reasons, the insurgency movement is gaining momentum.

He discussed and demonstrated a method called turbanology to help illustrate the importance of understanding cultural information. In Afghanistan, the kind of head covering you wear often tells a lot about a person – even from a distance. This is the kind of information HTTs, PRTs, and soldiers can use to gain a better understanding of their environment and to achieve mission success.

QUESTION AND ANSWER SESSION:

In response to a question about the emergence of Pashtu suicide bombers, Thomas Johnson responded that it is a new phenomenon that was grafted onto the Afghan insurgency from the Iraqi experience. Until 18 months ago, many suicide bombers were outsourced to disgruntled Punjabis and Arabs. The Pashtu bombers today are young men who never experienced traditional upbringing in a tribe. They lack the cultural norms that would keep them away from suicide bombings. Al Qaeda is taking advantage of this population.

Nick Pultorak asked whether there was a limit to hospitality that an Afghan would give a guest.

Thomas Johnson responded that there is no way to do this in rural areas and Afghanistan is 80 percent rural. The US Seals experienced this. The villagers gave them hospitality at the expense of their own safety. It is so foreign to us that we do not understand it. The Pashtu believe they raise their people better than the west. They believe their judicial system works better than ours. The Pashtu's have practiced pure democracy for

millennia and see no reason to follow our system. Imposing our government on them is the wrong thing to do.

Brian Meadows asked how we would get this information to the field.

Thomas Johnson said that he is in regular contact with people in field. He is not sure that they are not getting the data. The answer to the question is what the HTS is all about.

Elisa Bienenstock stated that when talking about the complexity of many socio-cultural tools; compare socio-cultural models to the radar. When the radar was first introduced, it took a PhD to read the output. Today anyone can do it. It is a gradual process. There is a culture that needs to develop; it does not happen overnight.

SESSION 4: SUPPORTING ANALYTICAL DISCIPLINES

Bob Popp, National Security Innovations

Bob Popp, NSI, introduced session four. The focus of this session is on different analytic approaches to support what was heard in first three sessions. Laurie Fenstermacher, AFRL, will present first a recap of the social science modeling and info visualization workshop held in late January 2008. On the next panel, Claudio Cioffi Revilla, GMU, will talk about his work. Elisa Bienenstock, NSI, will discuss quantitative sociology. Jim Llinas from Suny Buffalo will speak about data fusion. Ted Senator, SAIC, will talk about data mining. Gary Jackson will talk about applied behavior analysis techniques.

PANEL 4.1: MODELING AND VISUALIZATION WORKSHOP SUMMARY

Laurie Fenstermacher, Air Force Research Laboratory

LAURIE FENSTERMACHER

Air Force Research Laboratory (AFRL)

The objective of the Social Science Modeling and Visualization Workshop, held in January 2008, was to provide a "mixing bowl" for social scientists, modelers, researchers and government stakeholders to discuss the state-of-the-art in methods/models/visualization and potential application in SMA efforts. The findings of the workshop are summarized in a workshop report, available from Bob Popp.

There are many benefits of modeling. Models can clarify complex situations, test assumptions, aid decision making, explore co-evolutionary motivations, and bound the expected and outlier behaviors. Modeling can capture dynamical, non-linear processes. They can be multi-level, multi-scale behavior representations. Models can address weaknesses in mental models by making assumptions explicit, helping to calculate consequences of assumptions and account for many factors, providing repeatability and controllability. Social science modeling has great potential as long as it is well documented, not overly complex, transparent, not over-sold and not over-promised. Models can provide a formal mechanism to attach evidence to the social science motivated models (relative likelihoods of outcomes) and/or enable the human-in-the-loop to link and assess evidence.

Other benefits include that models can assist in training decision makers on the consequences of their policies/decisions (e.g., virtual environments). Typologies help users orient to a society/culture, ensure relevant information is captured, expose key (and often not obvious) relationships, guide resource management/data collection, create ontologies for computational models. Models enable the exploration of

data and "what-if" analyses (e.g., alteration of the parts of a social network). Connecting (theoretical foundation of) social science modeling and (analytical tools of) system engineering enables predictions of likely behavior of actors based on government policies.

In relation to terrorism research in particular, modeling can achieve a number of objectives. Modeling provides non-intuitive insights into necessary factors, e.g., importance of organizational size, religious ideology, organizational connections, and control of territory in forecasting groups who are prolific killers. Modeling is useful for establishing grounded expectations about terrorism/terrorist actions, identifying trends/changes in trends, specifying indicators associated with terrorism/terrorist activity. Modeling can provide insights into a variety of issues: group dissident dynamics, impact of state repression/government policies/counterterrorism measures on violent/non-violent dissident groups, terrorist tactics, "when" and "why" groups make the transition to become very violent. Modeling can also forecast the most probable response of a given group in a situation and identify actions that maximize the probability of a desired response.

In regard to WMD insights, modeling is also useful. Modeling provides insights into key relationships and effects (e.g., possession of nuclear weapons does not increase/decrease disputes; it increases diplomatic leverage/influence). Modeling helps to identify necessary variables/factors for pursuit of chemical, biological, radiological, and nuclear (CBRN) terrorism, but data is very sparse. Models can focus collection on indicators in terrorist environment related to potential use of WMD (e.g., training materials, manifestos, statements).

Visualization is a critical component of modeling. Therefore, it needs to be considered from the beginning of a project. It is effective when analytic tools cannot provide the answer alone -- as model complexity increases, the need for advanced visualization and analysis also increases. Visualization (e.g., Visual Analytics) enables keeping the analyst and decision makers "in the loop", helping them interact with the data and fostering ownership of modeling/analysis result. It enables analysts to view and interpret multiple kinds of data and to effectively sift through data. It helps analysts better understand their data -- much can be gained by visually inspecting data versus looking at a set of numbers or reading text. Visualization "provides the user with layers and multiple ways to look at the data to draw conclusions, relations and observations". Finally, it can reveal the structure of an event (e.g., pandemic outbreak), critical locations and enable early detection and efficient targeting.

Visualization serves a variety of functions include support of knowledge management, contextual visualization (e.g., system within a system-of-systems), relational visualization (e.g., words in a visual thesaurus), identification of key features/nodes/information, and understanding of decisions/complex processes, evaluation/inspection of the data (highlight "bad" data and outliers). Virtual worlds/gaming can enable research into the mechanisms of natural occurrences (e.g., plagues); gaming worlds can simulate large scale entities (economies, societies). Hardware and processing improvements enable improved rendering capability, detailed simulation of volume and surface phenomena/effects, interactive displays, traversal of multiple levels of data, massive data processing.

Although it seems obvious, machines should be allowed to do what they do best. They are fast, inexpensive, transparent, reproducible, and can use unbiased data coding. Its goal should be to optimize coding systems versus duplicate human coding performance. We need coding schemes for multiple languages that are teachable and transparent. Computers enable characterization of documents (e.g., positive/negative) can greatly assist with sorting documents ("triage for the analyst").

Data is another challenge for modeling and visualization. Modelers need current/relevant data (e.g., to assess resilience, need data after attack or raised threat level). This includes a variety of (direct and indirect) data sources: telephone and Internet panel surveys, interviews, open source/internet. Open source information is critically important to studying terrorism. Often the information needed is sub-national, but the data is not collected at this level.

Validation of models is important. Models must be verified, valid, and credible (credibility is achieved when the end users accept the model results as correct). SME's are an important part of the model validation

process – true experts have "good instincts". Finally, "Expert opinion and feedback from users whose lives have depended on results may be the best validation we will get…".

The theoretic framework of social science modeling and visualization still needs work. Opportunities and challenges include the need for better, scientifically grounded theories. Better theories yield clarified assumptions, bias and improved means for problem scoping; yield efficacious and efficient data gathering/manipulation; support new analytic methods and better visualization; and yield common language to interpret analysis and new visualization strategies. One problem is that the number of peer reviewed journal articles on WMD is low. There is a vital need for social theory to identify generative patterns and structures within social complexities. It is very difficult to model terrorist motivations due to lack of consensus on concept of WMD, lack of validity of old and new assumptions, role of Islam, etc.

Many issues remain in the social science modeling and visualization environment. This includes the:

- Ability to integrate human and physical factors into predictive models
- Ability to integrate models
- Ability to rapidly develop models/model setup
- Ability to "validate" models, inadequate evaluation
- Ability to precisely articulate a problem, imprecision of key terms (e.g., "culture")
- Necessity of using SME's, difficulty with robustness, repeatability, pedigree when using SME input
- Theoretical models not "grounded in data/evidence"
- Inappropriate use of models – "what are the assumptions?"
- Need to be flexible (e.g., typologies) to deal with problem space

The federal government is interested in social science, but struggling to determine how to support it. It will take 4-5 years to reach robust funding levels. There is a great need to leverage industry/academia because the "private sector provides unique insights unavailable from academia and public sector". Funders are a major factor in advancing the science of modeling. However, they need to fund long-term efforts. Additionally, efforts must be made to ensure agencies are not funding work that is repeating previous efforts.

PANEL 4.2: EVOLUTIONARY AGENT-BASED MODELING AND GAME THEORETIC SIMULATIONS

Claudio Cioffi-Revilla, GMU
Elisa Bienenstock, NSI

CLAUDIO CIOFFI-REVILLA

George Mason University (GMU)

Claudio Cioffi-Revilla, GMU, gave a brief overview of area agent based modeling (ABM) and evolutionary computation.

ABM is a very new methodology in social science. This kind of methodology in the life sciences is called individual base modeling. ABM is successful in the social science arena because as in the traditional approach to formal model, the goal is reducing the social landscape to a set of meaningful variables and specific math equations. ABM brings to social science the ability to begin with social landscape of entities in world. It can identify basic relationships between actors.

To create this kind of model, first the modeler must identify the ontology and methodology on which it is based. Once the landscape has been identified and described, then relevant attributes and operations of attributes come into play at the mathematical foundation of model. It is a way to formalize the representation of the social world. In sociology, demography is the most formal mathematical field of social science. ABMs

are very effective in this field. But in many other areas of social science, the purely mathematical approach has had its limitations because of great difficulty of going from a highly dimensional space to small dynamical equations.

An agent based model is a formal computational model or social simulation consisting of autonomous actors with attributes (human, social, cultural) and multiple interactive relations. Evolutionary computation refers to a class of computational approaches, including some ABM development strategies, inspired by the biological principle of change, fitness assessment, selection, and elimination. This is an experimental area. It has not been used much in the social sciences.

All models are developed or designed to answer a class of questions. ABMs are typically designed to answer questions that are intractable through earlier and more traditional mathematical or statistical methods, such as complex emergent patterns in large systems of agents.

ABM is especially valuable in situations in which there are lots of heterogeneous agents of various kinds. Populations are situated in physical environment. The beauty of ABM is that it can conduct experiments about shifting policies. All can be done in silicon because you cannot manipulate players in the real world. ABM is a specific kind of simulation. It provides experimental setting for finding patterns that are not very obvious.

In closing, ABM and evolutionary computation have a joint role to play in social sciences. Not all basic social sciences is needed or obtained. Evolutionary computation is one strategy to understand and discover patterns not yet found to patch up and compliment parts of modeling in simulation. It is very promising.

ELISA BIENENSTOCK

Chief Human Science Officer, National Security Innovations (NSI)

Elisa Bienenstock, NSI, spoke about game theoretic simulations. Game theory is a branch of applied mathematics that formally models strategic behavior. There are three major areas or levels of game theory:

1. 1 Person: Models or risk or decisions under uncertainty
2. 2 Person: Models of decisions when the outcome of a decision is contingent on the decision of another
3. N-Person: Model of decision when the outcome achieved is contingent on the behavior of multiple actors

There are two approaches to Game Theory: formal (analytical, mathematical) and evolutionary (computational).

Game Theory is an approach to aid in articulating and understanding important factors of conflicts / disputes / coalitions. It is a formalization of essential elements of assorted interactions. There are two primary ways that thinking in a game theoretic way is useful. First, game theory provides: "names" for assorted "games." It requires formalizing games by distilling essential features of the situation. Isolating the important characteristics of a situation is essential. It makes it possible to find similarities in what otherwise appear to be different situations. It also allows reduction of complex situation to a "symbol" which can quickly convey meaning. Second, game theory uses mathematical operations or "solution concepts" to provide insight into non-obvious processes to compare and contrast; generalize; determine resource distribution (optimal; equitable); and determine efficiency and stability of solutions etc.

Game theory is neither prescriptive nor descriptive; nor is it normative. It is a theoretical tool. Game theory does not provide an answer it provides a helpful way to frame the question to better understand the situation. If the assumptions are correct, game theory will provide an answer. Some assumptions of the theory include (bounded) rationality, complete information, and preference hierarchy for players (utility).

Game theory is not one thing: it is a set of related models. It is "a rigorous consistent superstructure into which separate models all nicely fit" (Shubik, 1984). It is also a "patchwork" theory: "a surprisingly large number of ingenious and insightful solution concepts for n-person cooperative game theory have been proposed by many different authors. Each solution addresses some particular aspect of societal rationality, which is the possible proposed or predicted behavior of rational individuals in mutual interaction" (Shubik, 1984). Some examples of common names of games include zero sum, prisoner's dilemma, and chicken game.

In the 1950's, game theory became very popular as a way to conceptualize the conflict between the US and the Soviet Union during the Cold War. The games were two person games (zero sum and involved escalation (nuclear arms race)); and n-person games (coalition building and NATO decision making). They were very tractable and amenable to analytical game theory.

Social dilemma can be conceptualized as prisoner dilemma games. In the tragedy of the commons, everybody wants to use the commons for their benefit. However, overuse depletes the commons. Everybody hopes others will not use the commons. When formalized so everyone has the same motivations this is very tractable, when multiple factors are considered computation is needed. Evolutionary element (the players who accumulate the most points are "fit" to survive or reproduce).

Relaxing the "formality" of game theory, but preserving the idea of modeling "strategy", Robert Axelrod conducted an n-person prisoner's dilemma experiment with real people. He submitted strategies for an iterative "round robin." The winning strategy was that of Anatol Rapoport and called Tit for Tat. He published the book "The Evolution of Cooperation" in 1984. This launched an interdisciplinary area. The prisoner's dilemma game has been shown to be "resolvable" thanks to "direct reciprocity."

Research has shown that indirect reciprocity, in large groups, does not generate cooperation. There are no repercussions for bad behavior and no incentive to behave. However, reputation can allow cooperation to prevail. A simulation to demonstrate that social networks as an instrument of reputation can produce cooperation.

In conclusion, it is important to remember that people are not agents. Game theory requires assumptions about behavior. It is a simplification: a model. It provides insight and understanding but people are more complex then models. Game theorists use experimental methods to test theory on small groups. New "gaming" technology may allow large scale games / experiments. This will provide a way to validate simulated agent based game theory models. However, one cannot escape that garbage put into any system will result in garbage coming out.

QUESTION AND ANSWER SESSION:

Hriar Cabayan stated that he was not sure if anyone saw the deterrence report written last year by Allison Astorino Cortois of NSI. In the report, it stated that the key to deterrence was that the adversary makes decisions based on his own world concept and his perception of what actions the other side would take. The blue side was also trying to deter and anticipate what the other side is doing. How can these ideas be folded into game theory?

Elisa Bienenstock replied that there are three levels of game theory. N person game theory is the basis of what Allison Astorino Cortois does. Her specialty is to map people's preference in real world on to that. Once you have utility function, you can crunch the numbers.

Claudio Cioffi-Revilla stated that in the original report, it concluded that deterrence was not a panacea for dealing with all kinds of adversaries. There are many situations where adversaries will not be deterred.

Bob Popp stated that much of the focus has been on non-state actors. He remarked on the debate that some people want to apply state based notions to non-state actors. Many people say you cannot do that and there has been lots of debate.

Claudio Cioffi-Revilla agreed that even applicable parts of deterrence have been questioned as to their relevance to non-state actors. For example, the type of motivation that a tribal leader has may include incentives and values that are beyond scope of what state authorities can do. When a leader knows that to remain in power he has to run redistributed economy, it is too much to give up to reorganize his government. There may be nothing in his repertoire to induce that type of behavior. Deterrence is one side. Compellence is the other side of equation.

Patti Morrissey stated that if you go back to Elisa Bienenstock's presentation, if people are compelled to act based on reputation, you can factor that in. But how do you factor in honor enhanced by dying in a logical model?

Elisa Bienenstock replied that if their value system does not value life or their own life, but values something above life, then that is their value system. We just put it in computer system. There may be things worse or better than life. This is still bounded rationality. Sometimes it is satisfysing and making sure you do not get the worst outcome.

Jim Llinas asked if you learn anything from the slope at which things change. What do you learn when there is convergence and some final condition rises?

Elisa Bienenstock stated that the question was too general to answer specifically. However, if you are looking at escalation, you can ask at what point incentives can turn the function. If you know how to articulate the problem in an equation, you can get a response. If you know it is the point of acceleration then you know how and where to crank it – how to raise the cost of war. You can then ask what parameters are most effective at the least cost.

A. Arash stated that analysts make models in their heads. They may ask what level of workforce requires what level of input. Using a pyramid drawing they can sketch you six layers starting with raw data, info, knowledge (add context), intelligence (add relevance), wisdom, and enlightenment. In the intelligence layer, analysts must make models.

PANEL 4.3: DATA FUSION/INTEGRATION AND DETECTING PATTERNS IN HETEROGENEOUS DATA SETS

Jim Llinas, State University of New York at Buffalo
Gary Jackson, SAIC
Ted Senator, SAIC

JIM LLINAS

Research Professor, Executive Director, Center for Multisource Information Fusion, State University of New York at Buffalo

Jim Llinas, State University of New York at Buffalo, spoke about the impacts to the data fusion process. Jim Llinas started his presentation saying that as an outsider, he attended the workshop to see how the intelligence and information fusion communities can work together to help solve human dimension problems.

Data fusion technologies encompass a variety of characteristics. It is historically a deductively-based inferencing / estimation process. Observations and estimates are random variables but optimized for minimum uncertainty. The process is adaptive in various ways. It enables observation management, process adaptation, and beneficial actions. Most applications are to mission-level timelines but sometimes data fusion is applied in a forensic context.

The data fusion process estimation levels are listed below. Level 0 refers to sub-object data association and estimation (e.g., pixel/signal level data association and characterization in imagery, before objects are discerned). Level 1 refers to object refinement (observation-to-track association, continuous state estimation (e.g. kinematics) and discrete state estimation (e.g. target type and ID) and prediction). Level 2 refers to situation refinement: object clustering and relational analysis, to include force structure and cross force relations, communications, physical context, etc. Level 3 refers to impact assessment ([Threat Refinement]: threat intent estimation, [event prediction], consequence prediction), susceptibility and vulnerability assessment. Level 4 refers to process refinement: adaptive search and processing (an element of resource management). Level 5 refers to human-fusion process interface.

The data fusion community works with real and dynamic states. Surveys refer to human order of battle information. There is an a priori assertion of the state of affairs as problem begins. Then the plan is thrown away and the dynamics of world begin to take shape.

The nature of the human dimension problem in regards to data fusion results in the following conditions.

- Limited Domain Knowledge
- Atypical Sensing/Observational Techniques—Hard (Electronic) and Soft (Human) Sensors
- Human-based Observational Sources Uncalibrated
- Extensive array of Contextual Information is pertinent
- Achieving the Emergent capability requires Runtime Inferencing/Estimation
    - Emergent-detection Capability ~ What Timelines?
    - Mission timelines not same as Forensic timelines
- Architecture: must fit into Net-Centric, Service-Oriented Framework
- Users: Embedded Human Terrain COI

In domain knowledge, a priori, deductive knowledge foundations for these problems is evolving and incomplete and uncertain. The impact results in a requirement for hybrid inferencing/estimation, requirement for knowledge model adaptation, the need for precursors/triggers of violent behavior, and disparate representations of uncertainty IAW depth of knowledge.

In reference to source characterization, conventional (hard/sensor data) fusion exploits knowledge of observational and estimation-based errors. To do source characterization for human observers observations can be couched in reliability of sources and uncertainty modeling of observational errors.

Data association is a core data fusion function. It determines which sets of observations are to be used for estimation-updating. It assesses the closeness of observations to each other or to a predicted entity state. It develops an association score for each feature/attribute contained in the observations. It is used for locational association.

Data fusion models can be used to analyze a broad set of areas. The physical world is the easiest to accomplish. This includes weapon systems and weather reports. The information domain is harder. This includes symbolic representation and interpretations of data and models. The cognitive world includes beliefs, values, and emotions.

TED SENATOR

Vice President, Chief Technology Officer, SAIC

Ted Senator, SAIC, spoke about data mining successes and challenges for complex human behaviors. The defense and intelligence communities are making great progress in the social sciences. Many agencies now realize the importance of social science approaches, in combination with more traditional approaches.

Some successes include anti-money laundering (FinCEN), NASD (now FINRA) market surveillance, and tribe finding at FINRA (Friedland & Jensen/U Mass). One common theme among these successes is finding and leveraging hidden relationships.

One way to elicit information was exemplified during the NASD project. "Trusted trader" interviews were conducted. They enlisted participants who may be disadvantaged by bad actors. They were represented formally to keep the lawyers away. The data was validated and they made high visibility cases. This approach depends on institution cooperation.

Screening can work on several levels. First, it can be used to raise the bar, which removes low-cost attacks and reduces the pool of bad actors. More complex attacks require more steps to plan and execute and more opportunities to detect. New attack techniques are hard to execute right the first time.

Often, people want to know why if we can go to the moon we cannot accurately model social science. It is because people are intelligent agents with feedback. They have multiple incentives, multiple allegiances, and multiple groups. Humans also exhibit adversarial behavior. They provide inaccurate and misleading information. Behavior is not invariant.

There is a hierarchy of challenges in data fusion. The first is registration and entity resolution. The second is group and network detection. The third is feature construction and recognition. The fourth is complex event detection. The fifth is that concepts in models are inherently fuzzy. There is a prediction versus risk relation. And risk brings additional scrutiny and preparation challenges.

There are many technical definitions of data mining. The three most relevant are described below. Fayyad et. al. posit that it is the non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. Jensen states that it is a process that uses algorithms to discover predictive patterns in datasets. Jonas and Harper state that it is the process of searching data for previously unknown patterns and often using these patterns to predict future outcomes.

The political definitions of data mining are even more numerous. One used by Senator Patrick Leahy is "the collection and monitoring of large volumes of sensitive personal data to identify patterns or relationships." Senator Feingold listed a much longer definition in March 2004 that takes up nearly three paragraphs, which is not listed here.

GARY JACKSON

Chief Scientist/Engineering Manager, Systems and Technology Division, Reconnaissance and Surveillance Operation, SAIC

Gary Jackson, SAIC, spoke about data fusion, integration, and detecting patterns in heterogeneous data. He spent 25 years working on automating processes. He has built a number of applications from the basic science approach including the automation of applied behavior analysis.

SAIC is working to create an automated multi-source, multi-sensor, multi-data fusion tool. One way they are doing this is by looking at textual information. They take documents of different types, use ABA tools, and extract antecedents. The tools can grab information, bring it together, and end up with predictive models.

Working toward multi-focused analysis, the tool must overcome the problem of sorting through tons of information. SAIC developed tool guided exploration across data types to help take the text corpus, pull out primary features, and guide it with documents. The process is now automated to the point where conversion of foreign and English language text to machine readable pattern classification code is possible.

The tool can be applied to tradecraft problems. The objective is the extraction of key adversary tradecraft and operating instructions from such sources as the Al Qaeda training manual, past incidents, HUMINT, recovered IED construction instructions, and analysis of past IED construction locations results in

behavioral profile attributes that may be layered with existing GIS location attributes to depict likely IED construction locations.

Commercial ABA based Checkmate system assesses threat in real-time (up to 1,000 entities tracked and assessed every 100 milliseconds) on an ongoing basis. It detects degree of deception and expertise to cause damage. It was validated independently by SAIC, Sandia, and Emagined Security (Berkeley).

The next steps including fusing cross area applications with code. New areas can be explored such as sentiment analysis.

QUESTION AND ANSWER SESSION:

Claudio Cioffi-Revilla asked Ted Senator why he built a model using purely inductive procedures.

Ted Senator replied that that is what data mining is – building models when you only have data. It is not necessarily the best way, but that is what data mining is.

One member of the audience asked Gary Jackson what variables he uses to predict terrorism looking a physical tradecraft environment.

Gary Jackson responded that if one is looking at terrorism, you take text accounts or media accounts. You look at examples of past terrorist attacks, pattern classifiers, and then validate blind cases. If you hit accuracy, you look at how to predict in future. If you train a certain way, you cannot change anything.

Elisa Bienenstock stated that she did work trying to classify different types of behaviors based on books on terrorism. However, if a tool is not trained to a specific area, they will not act as they are supposed to act.

PANEL 4.4: LARGE DATA SETS AND KNOWLEDGE EXTRACTION TECHNIQUES

Andy Wilson, Sandia National Laboratories
Philip Kegelmeyer, Sandia National Laboratories
Ben Mann, DARPA

ANDY WILSON

Sandia National Laboratories (SNL)

Andy Wilson, SNL, spoke about how large data and visualization efforts are becoming an ever growing avalanche.

Within large data and visualization efforts, there is no one true schema. Efforts must "adapt or die." Self-describing formats are good. They are primarily based on XML, SQL, and OO databases. Loading all the data is hopeless. Therefore, they try to load the structure of the data. This results in maximum expressiveness and minimum size. The structure of the data for documents includes terms, entities, and concepts. The structure for transaction, travel and communication records include people, places, organizations and relationships. However, smart loading still requires human guidance.

Listed below are guidelines for mastering scale. This is important when a database has hundreds of thousands of entities and millions of relationships. The first lesson is do not show the full graph unless you absolutely have to. It is too big to make any sense. Instead, show abstractions, clusters, patterns, groups. Also remember to look at outliers. The second lesson is that tools should let you ask questions and propose hypotheses. Think about it as working with a reference librarian instead of a card catalog.

Different views of the data show different properties. The "aha" moments happen when everything falls into place. This is easier when you can see all the pieces of the puzzle. You can view the model as an overlay, side-by-side, or small multiples. It also helps to see the interactions between views.

A sandbox is an environment for "what if" and for constructing arguments. It should have quick access to exploratory tools and answer the question, "How are these actors connected?" The system's job is to track workflow and allow for annotations and assumptions. The analyst can backtrack and replay, branch and recombine. He may also retract or change assumptions. The output should embody a complete chain of reasoning.

PHILIP KEGELMEYER

Sandia National Laboratories (SNL)

Philip Kegelmeyer, SNL, has been working for about a decade on ways to practically do predictive analysis on real world data. The problem is the immense amount of noisy data generated. Noise in this case does not only mean noise in the data collection, it means error in the human labeling of the training cases on which predictive analysis depends.

The scale and speed of the data meant that the classic, craftsman approach to machine learning is no longer possible. The answer is a commodity, hands-on model. This is a challenging task. Not only does commodity mean that we have to develop algorithms that are robust in the face of noise, skew, and the host of other ills; it also means that we have to work out how to set all algorithm parameters from the data or the context, so that the human analyst need not concern themselves with them.

There is no single algorithmic magic bullet. Work at Sandia has slowly established a set of principles, and algorithms and methods for implementing them. Since the nature of data has been changing, some of these principles break, or even reverse, long standing pattern recognition practice.

The specific sort of predictive analysis Sandia works with is "supervised pattern recognition", which means learning to label new data based on what you saw in old data. The main thing that distinguishes "supervised pattern recognition" from the unsupervised sort, from clustering and exploratory data mining, is the "supervision" provided by the existence of groundtruth data. That is, you start with training data that has attributes and labels. The point is that each sample has a label, probably from human groundtruthing, indicating whether or not the response was ideological. Each response also has a number of attributes, measurements, any descriptor that might be pertinent. The idea is that there is presumably, in reality, some unknown function that operates on the attributes to generate the labels, to predict whether a response will be relevant. A pattern recognition algorithm inspects the training data and builds an estimate of that function and then uses that estimate to map new, unlabeled data points to a predicted label.

There are lots of pattern recognition algorithms; some you have likely heard of are decision trees, support vector machines, neural nets, nearest neighbors, naive Bayes. Ensembles are the first and the most powerful of the techniques to know about, as they permit you to squeeze all possible accuracy out of your data.

Bozos are used because they are typically individually much less expert than the expert and because they are based on subsamples, tiny, tiny fractions of the original data. And yet, in aggregate, they are not only as accurate as an ensemble of experts, they can sometimes be more accurate. Ensembles exploit the power of multiple perspectives.

Ensembles are how to automatically get as much accuracy as possible out of your groundtruthed training data. And that is important, but generally it is not the sole problem. Generating the groundtruth in the first place tends to be tedious and expensive, because it usually involves human time and judgment (and error). As Dr. Stone said yesterday, she gets "boredom errors" out of the people who groundtruth her text data. Similarly, we also heard at length from Dr. Pate how on the difficulties of coding her Minority at Risk data.

To deal with underrepresented data, you can synthesize training samples, using a specific technique called SMOTE. That is, if you do not have enough of the right data, you just make more. This sounds scary, but it is probably much better than just replicating or oversampling the known true samples from your minority class. The basic idea that you look at the parts of feature space that are closer to your minority class than any other class, and select random unpopulated spots from that space to be additional minority points. Mathematically, the existing minority points are used to build a non-parametric estimate of the underlying distribution of the minority class, and then we simply sample from that distribution.

By construction, the new points generated by SMOTE come prelabeled. Labeling the points is expensive, but some actions may be taken to mitigate the cost. One way is to label only a few of the points, as many as you have the time, money, or energy for, and then let the system bootstrap itself. This works because predictive analysis can assign not only a label to unlabeled data, but also a confidence in that label. So the idea is that you take what little groundtruth you have, build an ensemble, label all the unlabeled data, and look at the confidences. Anything that is really confident, say 95 percent or above, you accept and treat as if had been human-generated ground truth. With this extra ground in hand, you build a new ensemble, and repeat. There is no provability around this, but it often works quite well.

Another approach, one that is complementary to semi-supervised and so can be combined with it, is active learning. The idea here is that you do not have the human groundtruth data in random or serial order. Rather, it is possible to do some upfront analysis, and figure out which of the unlabeled samples would be most useful to label, the ones that would shed the most light with the most confidence on the rest of the unlabeled data.

BEN MANN

DARPA

Ben Mann, DARPA, spoke about his effort to find fundamental laws in biology. If you want to understand something, you want to understand its principles.

The real world has a crazy shape to it. If you put a crazy shape in a big shape, scatter points across shape, and throw the shape away, what do you have? These questions and others comprise the foundation of Ben Mann's work. He asks whether he could construct a scaffold for the data that can say something about the dataset. Can one construct geometric objects that have much more information than numbers? The null hypothesis is random.

It turns out that there is an FDA sanctioned test called APACHE used in hospital Intensive Care Units. It represents a lot of discrete data that is easy to get, but hard to figure out its pattern. If one patient has higher APACHE score than the other – then that patient is more likely to live than the other. If you have a database of 150,000 patients with discrete data (heart rate, blood pressure) and the outcome you find that despite the APACHE score, the error rate is 50 percent. DARPA built a better analysis model with 20 percent error rate.

The DARPA model looks for persistent clustering and for high dimensional features. The geometry in the data set might help pull out information. Clustering algorithms are related to naturality. When I make construction, I have some kind of mathematical object with function to another math object. A natural question is "if I make a computation over here and move it over here, does it inform making inferences about cluster over there?" The answer is no. If there is a choice of scale involved it may pick out what you really want.

The idea is that there are many parameters. DARPA is trying to cluster or understand how systems work under various parameters. Use idea of naturality to compare equations as you go along. The idea of algorithm is going in and having lots of parameters they can vary to build geometric models to pull features out. This method has been used successfully in biology.

If you look at physics, the math needed to model it is very tame math. In biology, the math needed is fundamentally different. DARPA is trying to invent the math that goes with biology. Some of the math might be relevant to applications in economics or social science. Instead of thinking about numbers, point clouds and data are fundamental objects.

Elisa Bienenstock stated that even though we have not discovered the laws of social science, there are some regularities. They have not been captured in math.

Ben Mann replied that he believes that some ideas in string theory might also apply. Physicists believe that math accurately predicts how physical objects behave dynamically. Biologists tend not to believe in these laws. If you take math for physics and try to do that, they are correct.

AGENDA

Day 1	24 April 2008	
Time	Topic	Speakers/Panelists
0830-1020	Session 1: The Evolution of Violent Extremism	Clark McCauley- Bryn Mawr (START) Todd Helmus – RAND Amy Pate – UMD (START)
1040-1150	Panel 1: Government-Sponsored Research on Violent Extremism	Allison Smith - DHS Matt VanKonynenburg - DoD Shawn VanSlyke - FBI LtCol Lyons - MCIA
1150-1300	Working Lunch	
1300-1440	Session 2: Understanding Socio-Cultural Dynamics- Tradecraft and Observation from a distance	Kelcy Allwein - DIA Matt Page - ODNI Joshua Sinai – Analysis Corp Bill Diggins- Gallup Rebecca Goolsby – ONR
1500-1545	Panel 2.1: Socio-Cultural Tradecraft and Open Source Collection Requirements Management	Lizbeth Sydnor – DIA Swanson - Delphi Intl
1545-1630	Panel 2.2: Remote observation using large data sets	Hugh McElrath –PNNL Robalyn Stone - SSA
Day 2	**25 April 2008**	
0800-0805	Session 3: Understanding the Human Terrain – Tradecraft and Streetcraft Methods	Patti Morrissey - OSD Kathleen Kiernan – Kiernan Group
0805-0930	Panel 3.1: Applying Law Enforcement Methods to gathering Human Terrain	Gary Greco - DIA Michael McNicholas – Phoenix Group Jim Powlen – Logos Tech Rich Shimon – Dept of Commerce
1000-1045	Panel 3.2: Applying Social Science methods to understanding the Human Terrain	Stefan Kaszubowski - YouGov Andrea Jackson - IDA
1045-1130	Panel 3.3: Soldiers as Human Terrain Sensors	LTC Lynda Granfield - DoS MAJ Craig Gendreau - Army
1130-1300	Working Lunch- Civilians as Human Terrain Sensors	Thomas Johnson - NPS
1300-1330	Session 4: Supporting Analytical Disciplines	
1330-1400	Panel 4.1: Mod/Viz Workshop Summary	Bob Popp - NSI Laurie Fenstermacher – AFRL
1400-1500	Panel 4.2: Evolutionary Agent-based modeling and Game Theoretic simulations	Claudio Cioffi-Revilla - GMU Elisa Bienenstock - NSI
1500-1600	Panel 4.3: Data Fusion/Integration and Detecting Patterns in Heterogeneous Data Sets	Jim Llinas – Univ of Buffalo Gary Jackson - SAIC Ted Senator - SAIC
1600-1700	Panel 4.4: Large Data Sets and Knowledge Extraction Techniques	Andy Wilson - SNL Philip Kegelmeyer - SNL Ben Mann – DARPA

ATTENDENCE LIST

Name	Org	Email	Telephone
Allwein, Kelcy	DIA	kelcy.allwein@dia.mil	202-231-5958
Arash, A.	OSC	arasha@rccb.osis.gov	703-613-5231
Benson, JD	AFRL		937-255-9945
Bienenstock, Elisa	NSI	ejb@natlsec.com	818-625-9047
Cabayan, Hriar	DDR&E	hriar.cabayan@osd.mil	703-746-1453
Canna, Sarah	NSI	scanna@natlsec.com	301-466-2265
Chesser, Nancy	SMA Team	nancy.chesser@js.pentagon.mil	703-697-0211
Cioffi-Revilla, Claudio	GMU	ccioffi@gmu.edu	703-993-1402
Connors, Frank	DTRA	frank.connors@dtra.mil	703 767-3355
Cross, Glenn	FBI	glenn.cross@ic.fbi.gov	202-324-3824
DelVecchio, Justin	CUBRC	delvecchio@cubrc.org	716-204-5139
Diggins, Bill	Gallup World Poll	Bill_Diggins@gallup.com	609-279-2223
Fenstermacher, Laurie	AFRL	laurie.fenstermacher@wpafb.af.mil	937-255-0879
Gendreau, Craig	95th CIV Affairs BDE	craig.warren.gendreau@us.army.mil	910-396-0685
Giever, Dennis	Indiana U of Penn	dgiever@iup.edu	724-388-9709
Goolsby, Rebecca	ONR	goolsbr@onr.navy.mil	703-588-0558
Granfield, Lynda	DoS	GranfieldLM@state.gov	202-203-7779
Greco, Gary	DIA, JITF-CT	gary.greco@dia.mil	
Gyurko, Angela	DOE	Angela.Gyurko@in.doe.gov	202-586-8017
Helmus, Todd	RAND	helmus@rand.org	
Henze, Roger	GISC	henzerh@stratcom.mil	402-232-2860
Hunt, Carl	IDA	chunt@ida.org	703-933-6532
Jackson, Andrea	IDA	ajacks13@jhu.edu	202-714-8278
Jackson, Gary	SAIC	gary.m.jackson@saic.com	
Johnson, Alexis	SAINC	amjohnson@sainc.com	
Johnson, Thomas	NPS	thjohnso@nps.edu	
Kaszubowski, Stefan	YouGov	stefan.kaszubowski@yougov.com	00447971540316
Kaupas, Ryan	NGA	rkaupas@sciasolutions.com	202-290-9859
Kegelmeyer, Philip	Sandia	wpk@sandia.gov	925-294-3016
Kelly, Michelle	NCTC	micheldk@nctc.gov	571-280-6699
Kiernan, Kathleen	DDR&E	kiernangroup1@comcast.net	904-303-0243
Kowal, Dennis	IDA	dkowal@ida.org	
Kusiak, Pauline	OUSDP	pauline.kusiak@osd.mil	703-692-3905
Kuznar, Larry	NSI	lkuznar@natlsec.com	260-636-1041
Lark, Rosemary	Homeland Security Inst.	rosemary.lark@his.dhs.gov	703-416-8493
Laurence, Jonathan	BC	jonathanlaurence@aol.com	617-230-0387
Llinas, James	U. Buffalo	llinas@eng.buffalo.edu	716-645-2357 ext 2105
Lyons, LtCol	MCIA	tlyons@mcia.osis.gov	

Name	Org	Email	Telephone
Mann, Ben	DARPA	benjamin.mann@darpa.mil	
McCauley, Clark	Bryn Mawr (START)	cmccaule@brynmawr.edu	610-526-5017
McCormick, Nahid	OSC	nahidm@rccb.osis.gov	703-613-5795
McElrath, Hugh	PNNL	hugh.mcelrath@pnl.gov	703-769-9510
McNicholas, Michael	Phoenix Group	mmcnicholas@phoenixgrouppanama.com	
Meadows, Brian	SPAWAR	bmeadows@spawar.navy.mil	619-553-2823
Morrissey, Patti	USD(I)	Patricia.morrissey.ctr@osd.mil	703-602-5870 EXT 7122
Numrich, S K	IDA	snumrich@ida.org	703-845-6807
O'Connor, Jennifer	DHS	jennifer.o'connor@dhs.gov	202-254-6716
Page, Matthew	ODNI	matthetp@dni.gov	703-482-9015
Pate, Amy	UMD (START)	apate@start.umd.edu	
Peters, Dave	DoC	DPETERS@bis.doc.gov	202-482-6446
Peterson, Christine	DIA/DIOSPO	Christine.Peterson@dia.mil	202-231-0206
Popp, Robert	National Security	rpopp@natlsec.com	781-864-1347
Powlen, Jim	Logos Tech	lpowlen@logostech.net	703-584-5696
Pultorak, Nick	USSOCOM/J2	nicholas.pultorak.ctr@socom.mil	813-826-1639
Rhem, Sam	SMA Team	samuel.rhem@js.pentagon.mil	703-692-0160
Ryder, Dorene	BBN	dryder@bbn.com	703-284-1347
Senator, Ted	SAIC	ted.e.senator@saic.com	703-469-3422
Shannahan, Lt Col Mike	STRATCOM	shannahm@stratcom.mil	402-232-2854
Shimon,Rich	DoC/OEE	rshimon@bis.doc.gov	703-487-9300
Sinai, Joshua	The Analysis Corp.	jsinai@theanalysiscorp.com	703-744-1523
Smarick, Kathie	START	kjsmarick@start.umd.edu	301-405-6739
Smith, Allison	DHS	allison.smith@dhs.gov	202-254-2357
Smith, Teresita	DIA OS Manager	teresita.smith@dia.mil	202-231-5413
Speed, Ann	SNL	aespeed@sandia.gov	505-845-0143
Stites, Frances	DHS	francis.stites@associates.dhs.gov	202-254-6038
Stone, Robalyn	SSA	robalyn@socialscience.net	614-527-5200 EXT 113
Strong, Timothy	US Army - Ft. Bragg	timothy.strong3@us.army.mil	910-391-8547
Swanson, Scott	Delphi Intl.	scott.swanson7@us.army.mil	312-659-3000
Sydnor, Lizbeth	DIA	lizbeth.sydnor@dia.mil	202-231-8143
Usher, Bill	RRTO	billusher77@verizon.net	703-742-9279
VanKonynenburg, Matt	DoD	matthew.vankonynenburg@dia.mil	
VanSlyke, Shawn	FBI	SVanSlyke@FBIAcademy.edu	
Veazie, CAPT Todd	JS	todd.veazie@js.pentagon.mil	703-571-1892
Wagner, John	SNL	jswagne@sandia.gov	505-845-7295
Wall, Joel	DHS S&T		202-254-5418
Wilson, Andy	Sandia	atwilso@sandia.gov	505-844-1089
Yaphe, Judith	NDU	Yaphej@ndu.edu	202-685-2224