# VACCINE ANNUAL REPORT – YEAR 2
## April 1, 2010 – March 31, 2011
Cooperative Agreement No. 2009-ST-061-CI0001

PURDUE UNIVERSITY

HOMELAND SECURITY UNIVERSITY PROGRAMS
TODAY'S RESEARCH & EDUCATION, TOMORROW'S SECURITY

# Table of Contents

# I. Summary

## A. Overview

The Visual Analytics for Command, Control, and Interoperability Environments Center (VACCINE) was established on July 1, 2009 to co-lead, with Rutgers University, the Department of Homeland Security's (DHS) Center of Excellence in Command, Control and Interoperability. VACCINE's mission is dedicated to creating methods and tools to analyze and manage vast amounts of information for all mission areas of homeland security. VACCINE accomplishes its mission through an integrated program of research, education and outreach, spanning the disciplines of visualization and computer graphics, engineering, computer science, geographic information systems, cognitive psychology, information technology, and emergency management and public safety. VACCINE is an international center with overall management and the lead research component handled by Purdue University and the education/MSI mission directed by the University of Houston-Downtown with Georgia Tech providing an advisory role. During early 2011, Morgan State University was added as an MSI partner, bringing our MSI participation to 5 schools. The VACCINE team is currently comprised of the following 19 universities with the associated Principal Investigator at each school:

| *University* | *PI* |
|---|---|
| Purdue University | Dr. David Ebert |
| Georgia Institute of Technology | Dr. John Stasko |
| Pennsylvania State University | Dr. Alan MacEachren |
| Stanford University | Dr. Pat Hanrahan |
| University of North Carolina at Charlotte | Dr. William Ribarsky |
| University of Washington | Dr. Mark Haselkorn |
| Florida International University (MSI) | Dr. Shu-Ching Chen |
| Indiana University | Dr. Shaun Grannis |
| Jackson State University (MSI) | Dr. Loretta Moore |
| Morgan State University (MSI) | Dr. Timothy Akers |
| University of Texas at Austin | Dr. Kelly Gaither |
| Virginia Tech | Dr. Chris North |
| University of Houston, Downtown (MSI) | Dr. Richard Alo |
| Simon Fraser University, CA | Dr. Brian Fisher |
| University of British Columbia, CA | Dr. John Dill |
| University of Stuttgart, Germany | Dr. Tom Ertl |
| Justice Institute of British Columbia, CA | Mr. Jack McGee |
| Navajo Technical College (MSI) | Prof. Mark Trebian |
| Swansea University, Wales | Dr. Min Chen |

Figure 1

VACCINE is strategically positioned to support the Department of Homeland Security in confronting the challenges of safeguarding our nation in preventing, responding to, and

recovering from events in applications including criminal investigation, health surveillance, fraud detection, emergency response, natural disasters, and border and infrastructure security. VACCINE focuses on the research, development, and deployment of interactive visual analytic environments for communicating and disseminating information and deriving insight from the massive homeland security data deluge. VACCINE is developing tools to help homeland security personnel, responders, and decision makers make sense of the sea of text, sensor, audio, and video data by developing powerful analytical tools and interactive visual decision making environments that enable quick, effective decisions as well as effective action and response based on available resources. VACCINE integrates data and analysis into interactive visual displays to enable users to make discoveries, decisions, and plan action. Applications include public safety, public health, and emergency response. Turning massive data into actionable knowledge through visual analytic techniques is vital to the mission of the Infrastructure Protection & Disaster Management Division of The Department of Homeland Security, as well as all of the mission areas of DHS.
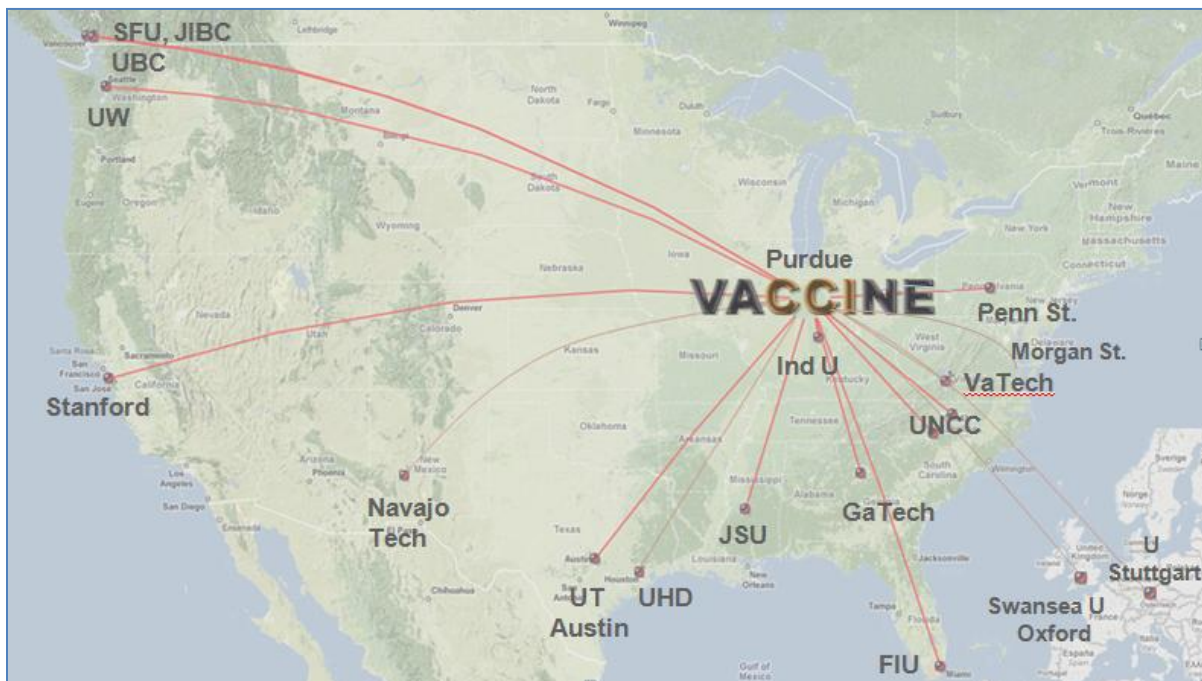


Figure 2: VACCINE international team members and partners.

**B. Accomplishments**

During year 2, the VACCINE team worked diligently to create a robust portfolio of research and academic projects as well as make significant progress in our outreach activities. Highlights of our team's efforts include the data in Figure 3 as well as these activities:

- VACCINE continued building our multi-agency public safety consortium. Memorandums are being negotiated with agencies outside Indiana.

- Our Visual Analytic Law Enforcement Toolkit (VALET) has now been deployed at the Lafayette, West Lafayette, and Purdue Police Departments. We are working with the area crime analyst, Steven Hawthorne, to test, evaluate and provide feedback.

- Jigsaw has been deployed at the Indianapolis Police Department, the West Lafayette Police Department and the Lafayette Police Department. We are working with the area crime analyst, Steven Hawthorne, to test, evaluate and provide feedback.

- VACCINE established a unique corporate affiliate program in the form of a non-profit Limited Liability Corporation, VACCINE, LLC. Membership is by invitation only. Corporations will be complementary rather than competitive. Motorola Solutions, Inc. became the inaugural donor to VACCINE, LLC with a gift of $40,000.

- We are working with the Indianapolis Gang Task Force in the development of the Gang Graffiti Automatic Recognition and Interpretation technology to bring them a tool that can translate and interpret gang graffiti for law enforcement.

- The project team at Penn State has been working actively with personnel in the Harrisburg Bureau of Police to transition the CrimeViz web-mapping software for their use. Mayor Linda Thompson publicly listed CrimeViz as one of the key developments towards improving public safety in Harrisburg.

- Developed and deployed the Coast Guard Search And Rescue Visual Analytics (cgSARVA) technology to the Operations and Planning division of the USCG as well as to District 9 Headquarters. The cgSARVA technology provides a detailed visual analytics model that now allows a thorough assessment and analysis of the Coast Guard's Search and Rescue mission on the Great Lakes.

- Partnering with CREATE, developed the PROTECT system for USCG security patrols being used in Boston, with expansion to other areas in Fall 2011.

- Our mobile phone hazardous material app has been deployed to Purdue police and is being deployed to personnel at the United States National Fire Academy (DHS/FEMA). Our tool allows a first responder to take a picture of the placard on a truck containing hazardous materials. It identifies the materials and provides response instructions quickly and simply. The information provided is tailored to the expertise level of the user, such as a police officer with minimal training (e.g., stand back 500 ft and call personnel with hazmat suits) to a fire fighter hazmat crew (e.g., treatment material, remediation protocols).

- VACCINE and CCICADA have been collaboratively working on projects in four research themes (Public Safety Coalition Projects, Enterprise Resiliency Environments, Event Evacuation, and International Collaborations) and two Education Themes (K-12 Programs and Professional Training).

- We are developing law enforcement tools and evacuation planning tools for the Charlotte Mecklenburg Police Departments.

- Expanded public health surveillance activities from Indiana to Pakistan with funding from the Nunn-Lugar act

- We held visual analytics workshops for high school teachers in 2010. Course modules were developed and are now in use in middle school and high schools in Indiana.

- An iPad application based on Florida International's "Integration Framework for Enhancing Emergency Response" system has been developed and presented to the personnel at Miami-Dade Emergency Management (MDEM), where an evaluation

activity regarding the developed application was successfully conducted. MDEM has expressed interest for us to further develop the system into an operational pilot.

- We are developing visual analytics flooding and preparedness tools for the DHS Natural Disasters, Coastal Infrastructure and Emergency Management Center of Excellence at the University of North Carolina.

- We continued developing our network of HBCU's, Hispanic Serving Institutions and Tribal Colleges and Universities for engaging in visual analytics and network security research and education. In early 2011, Morgan State University was added as a partner.

- In September 2010, the Justice Institute of British Columbia hosted a workshop centered on "visual analytics" in command and control interoperability environments for first responders. This workshop began to define the requirements that will lead to developing academic programs and relevant technology utilizing visual analytics in the public safety environment.

- VACCINE is finalizing plans for a joint United States-Germany research project to apply visual analytics to disaster prevention and crisis response, with a focus on critical infrastructures in logistics, transportation, food safety, digital networks and power grids at the national levels. This project is in collaboration with several international partners, NVAC, and NCFPD.

| VACCINE's Accomplishments (April 1, 2010 – March 31, 2011) | |
|---|---|
| Research Themes | 6 |
| Education Themes | 4 |
| Seed Projects Funded | 4 |
| Peer-reviewed/Refereed Journal Articles | 52 |
| Conference Publications & Presentations | 63 |
| Other Reports | 11 |
| Scholarly Journal Citations of Published Reports | 50 |
| Student Theses | 5 |
| Patents | 3 |
| Graduate Students Participating in VACCINE Programs | 40 |
| Minority Students Participating in VACCINE Programs | 9 |
| HS-STEM Student Theses | 5 |
| HS-STEM Students Enrolled | 12 |
| K-12 Student Participation | 120 |
| VAST Challenge 2010 Awards | 5 |
| Period 2 Base Funding | $1,684,990 |
| Additional DHS Funding (BOA, Supplemental, etc.) | $3,346,653 |
| Funding from other Agencies Leveraged | $9,088,825 |
| VACCINE, Limited Liability Corp. Funding | $40,000 |

Figure 3: Summary of VACCINE Accomplishments

## C. Business Model and Management Scope

The role of developing detailed operational requirements is critical for the successful transition of technologies into the various homeland security fields. The VACCINE Center of Excellence recognizes the need for significant end-user-in-the-loop involvement in the entire life cycle development process. In 2010, VACCINE refined our business model (figure 4) with the goal of connecting homeland security researchers with the first responder community. Understanding the requirements of our end-users is key to the research and development of technologies that can be transitioned to the broader public safety community. Our Public Safety Consortium consists of a number of law enforcement, fire, and emergency management agencies that allow data to be shared and research conducted furthering the development of next generation homeland security technologies.

The Consortium provides a framework where VACCINE can deploy its tools to the various agency members and our end-users provide in-field or real world evaluation feedback on their requirements. The group meets regularly to discuss ongoing requirements to technology being developed, improvements to technologies that have been deployed and are being evaluated, as well as to brainstorm tools that are needed. VACCINE sees this Consortium as a resource that can be utilized by not only VACCINE but also DHS S&T. The Public Safety Consortium continues to grow with expansion outside the State of Indiana anticipated for 2011. Our ultimate goal is to develop technologies that have been vetted and verified by having end-users in the development process. Technologies that have been designed in this fashion will be able to transition across a larger community of end-users geographically.



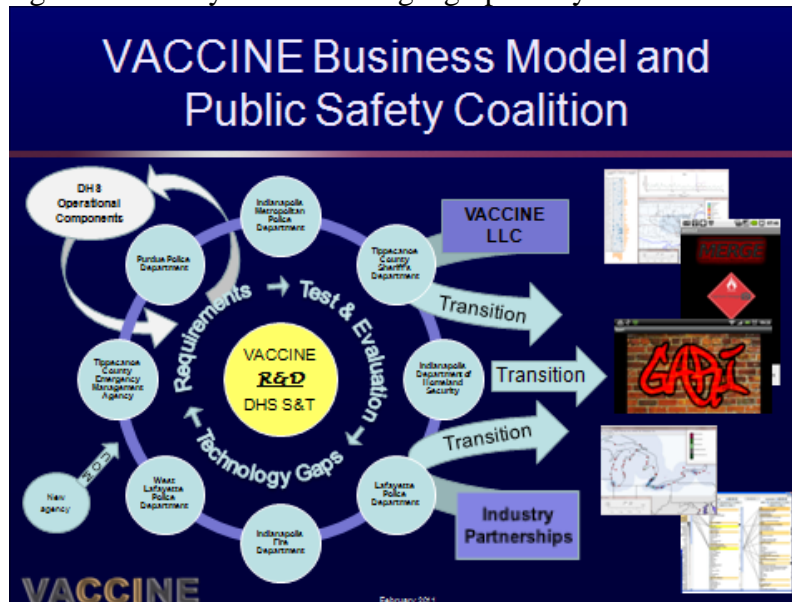Figure 4: VACCINE Business Model

Our team has a strong collaboration history and experience that is essential to foster internal collaboration, as well as collaboration with DHS CCI, NVAC, other DHS CoEs, government agencies, field personnel, and other researchers. Our Center's management organization has three components: management staff, a leadership board, and a senior external advisory board.

The management staff is intentionally lean in order to maximize research and education dollars while helping to minimize overhead. This team is responsible for overseeing the strategy, as well as the day to day operations of VACCINE, including interfacing with DHS, CCICADA and all external partners.

**The management staff** consists of the following:

- Center Director and Research Director – David Ebert
  David Ebert is a Professor in the School of Electrical and Computer Engineering at Purdue University, a University Faculty Scholar, a Fellow of the IEEE, and Director of the Visual Analytics for Command Control and Interoperability Center (VACCINE), the Visualization Science team of the Department of Homeland Security's Command Control and Interoperability Center of Excellence. Dr. Ebert performs research in novel visualization techniques, visual analytics, volume rendering, information visualization, perceptually-based visualization, illustrative visualization, mobile graphics and visualization, and procedural abstraction of complex, massive data. Ebert has been very active in the visualization community, teaching courses, presenting papers, co-chairing many conference program committees, serving on the ACM SIGGRAPH Executive Committee, serving as Editor in Chief of IEEE Transactions on Visualization and Computer Graphics, serving as a member of the IEEE Computer Society's Publications Board, serving on the IEEE Computer Society Board of Governors, and successfully managing a large program of external funding to develop more effective methods for visually communicating information.

- Director of Education and MSI Programs – Richard Alo
  In 2010, Dr. Alo became the Education Director for VACCINE. In this capacity, he oversees and directs the education strategy. Dr. Alo also oversees the integration of minority serving institutions into VACCINE research and education initiatives. Dr. Alo has served on several advisory and executive boards, including: Mathematical Association of America SUMMA Committee (Support for Undergraduate Minority Mathematics Awareness), Association of Departments of Computer and Information Science and Engineering at Minority Institutions (founding member), Coalition for Diversity in Computing (Chair), National Science Board NSF GRPA Committee, Multi Sector Crises Management Center (Board of Directors), and Houston Independent School District. Dr. Alo was selected as the 2002 Educator of the Year by the Hispanic Engineer National Achievement Awards Conference.

- Education Advisor – John Stasko
  In order to focus more on teaching and research, Dr. John Stasko assumed an advisory role for VACCINE's education programs. John received his B.S. degree in Mathematics at Bucknell University in Lewisburg, Pennsylvania (1983) and Sc.M. and Ph.D. degrees in Computer Science at Brown University in Providence, Rhode Island (1985 and 1989). He joined the faculty here at Georgia Tech in 1989, and his primary research area is human-computer interaction.

- Managing Director – Timothy Collins
  Timothy Collins is the Managing Director of VACCINE. He oversees and implements all operational aspects of the Center including strategy, marketing, operations, finance, business development and human resources. He is an accomplished professional having a 25 year operations background in the public safety and security arena. He is a retired Indiana State Police Officer reaching the position of Region Commander and is a former Air Force Officer with a specialty in Command and Control. He is a graduate of the FBI National Academy and has an MBA from Purdue University. His business expertise includes program and project management, business development and strategic planning.

- Assistant Director of Education and Engagement – Marti Burns
  Marti Burns left VACCINE in early 2011. She was the Assistant Director of Education and Engagement. In this position, she oversaw educational initiatives ranging from K-12 programs through undergraduate and graduate level work, to professional training programs. She has worked with homeland security education for the past five years, leading an interdisciplinary team of faculty in the development and implementation of new curriculum for master's level courses as well as new courses for the K-12 arena.

- Center Coordinator and Administrative Assistant – Jenny Kelly
  Jenny Kelly serves as the information and communication manager for VACCINE. Her duties include planning and scheduling meetings and appointments; organizing and maintaining paper and electronic files; managing projects; conducting research; and disseminating information to faculty, staff, and external partners. She managed a large scale multi-disciplinary medical practice in Illinois before moving with her family to Indiana. She began work as a legal assistant in a local Lafayette firm specializing in intellectual property, business transactions, litigation, estate planning, and employment law. In 2007, Jenny joined the Purdue Research Foundation working in the Office of Technology Commercialization before transferring to Purdue University.
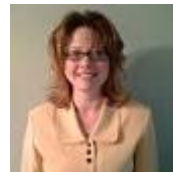
**The VACCINE Leadership Board** is responsible for the strategic decisions for the Center. This board is comprised of three permanent members plus three PIs from our universities on rotating two year terms, selected by the nineteen university PIs. The LB oversees the Seed Project review panel, advises on research, education, and outreach decisions and reviews and approves yearly Center plans.

- David Ebert (permanent)
- Timothy Collins (permanent)
- John Stasko (permanent)
- Pat Hanrahan (term)
- Kelly Gaither (term)
- Brian Fisher (term)
- Alan MacEachren (term)

- The VACCINE Senior External Advisory Board provides strategic advice, information, and recommendations to the Center on basic and applied research activities, educational issues, and activities and operations related to visual analytics and command, control and interoperability in the homeland security environment. This board consists of the world's top experts in visual analytics and homeland security from the scientific, business, and government communities. Current membership includes:

  - Don Brackman
    *President, National White Collar Crime Center*
  - Chief John Buckman
    *Past President, International Association of Fire Chiefs*
  - Dieter Fellner
    *Professor and Director, Technical University of Darmstadt & Fraunhofer Institute of Computer Graphics*
  - William Fox
    *Senior Vice-President, Global Anti-Money Laundering and Economic Sanctions Executive, Bank of America*
  - David Kasik
    *Enterprise Visualization Architect, Boeing*
  - Daniel Keim
    *Professor and Head, Information Visualization and Data Analysis Research Group, University of Konstanz, Germany*
  - Katherine A. Mitchell
    *Director STEM Initiatives, American Indian Higher Education Consortium*
  - Mike Mitchell
    *Director of Homeland Security Programs, Pacific Northwest National Laboratory*
  - Randall Murch
    *PhD, Associate Director, Center for Technology, Security and Policy and Adjunct Professor, Virginia Tech*
  - Haesun Park
    *Professor, Georgia Institute of Technology*
  - Andrew Vallerand
    *Director, Public Security Technical Program*
  - Leland Wilkinson
    *Executive Vice President of SYSTAT Software*
  - Tom Wyss
    *Indiana State Senator, DHS State and Local Officials Senior Advisory Committee for Homeland Security Advisory Council*

## II. Research and Development Scope

The overall goal of VACCINE Center research is to create innovative and effective integrated data and visual analytic environments—frameworks, methods, and software—that advance the start-of-the-art in analyzing massive, heterogeneous, incomplete, temporally evolving homeland security data for anticipating, detecting, and responding to homeland security mission needs. Therefore, VACCINE is developing new dynamic visual analytics techniques based on cognitive and perceptual principles that increase the effectiveness of the entire analysis process; create precision information environments; create an information discourse environment; and enable visual analysis, knowledge synthesis, and extraction of insight for actionable decision making, event detection, management, and response.

For effective visual analytics, we are tightly integrating computational processing, data management and visual representation to efficiently present relevant information and solution alternatives to the user and to improve both data management and statistical data modeling. To ensure successful deployment, scalability, and usability of developed technologies, we made a transition this year from our Mission Driven Research Project (MDRP) framework to a focus on six research themes that will tightly integrate with end-users from the beginning to the end of each specific project. Having our customers directly integrated with our projects allows them to provide input, project instigation, and relevant feedback and transition paths for DHS technology transfer. The specific projects are intended to be 12 to 36 months in duration with a clearly defined deliverable.

**Theme 1: Public Safety Coalition Projects** (state or local law enforcement, fire, emergency management, etc.)

**Theme 2: Federal Operating Component Projects** (TSA, FEMA, Secret Service, ICE, CBP, USCIS, Coast Guard)

**Theme 3: Enterprise Resiliency Environments**

**Theme 4: Event Evacuation**

**Theme 5: Visual Analytics for Security Applications**

**Theme 6: International Collaborations**

All of the projects associated with these themes have the following key pervasive homeland security problem characteristics: massive, time evolving, heterogeneous data of varying quality; multiple-scales of problem solving, planning, management, and action; coordinated information analysis; information security and privacy challenges; and a wide variety of needs, skills, and computing resources of involved personnel.

To achieve our goal of creating this visualization and analytics environment, our research plan is based on three foundational research areas that form our framework for addressing DHS mission directives: Interactive Visual Analytic Foundations, Visually-Adapted Analytical Techniques, and Investigative Analysis and Decision Making Environments.

For various logistical and timing reasons, Themes 3, 4, and 5 were significantly delayed in Year 2. It is anticipated that progress will be made on them in Year 3. The following is a short discussion on Themes 1 and 2.

**Theme 1: Public Safety Coalition Projects** (representative projects)

The role of developing detailed operational requirements is critical for the successful transition of technologies into the various homeland security fields. The VACCINE Center of Excellence recognizes the need for significant end-user-in-the-loop involvement in the entire life cycle development process. In 2010, VACCINE created a Public Safety Consortium with the goal of connecting homeland security researchers with the first responder community. Understanding the requirements of our end-users is key to the research and development of technologies that can be transitioned to the broader public safety community. The Consortium consists of a number of law enforcement, fire, and emergency management agencies that allow data to be shared and research conducted furthering the development of next generation homeland security technologies. The Consortium provides a framework where VACCINE can deploy its tools to the various agency members and our end-users provide in-field or real world evaluation feedback on their requirements. The group meets regularly to discuss ongoing requirements to technology being developed, improvements to technologies that have been deployed and are being evaluated, as well as to brainstorm tools that are needed. VACCINE sees this Consortium as a resource that can be utilized by not only VACCINE but also DHS S&T. The Public Safety Consortium continues to grow with expansion outside the State of Indiana anticipated for 2011.

In addition to the Indiana Public Safety Consortium, we have formed four public safety consortiums that will oversee the research and development process to guarantee the usability and utility of our products.

- Florida Division of Emergency Management's Region Seven Consortium
- UNCC Consortium - Charlotte Mecklenburg Police Department Consortium
- Washington State Consortium
- Penn State Law Enforcement Consortium

VACCINE's Indiana Public Safety Consortium began in 2010 with the Tippecanoe Co. Sheriff's Dept., Lafayette Police Dept., West Lafayette Police Dept., Tippecanoe Co. Emergency Management Agency, and Purdue University Police Dept. We have since added the West Lafayette Fire Department, Lafayette Fire Department, Indianapolis Metropolitan Police Department, Indianapolis Fire Department, and the Indianapolis Division of Homeland Security. VACCINE is currently negotiating Memorandums of Understanding with additional agencies in Ohio, Nebraska, and at the Indiana Intelligence Fusion Center. In 2011, we plan to expand the Consortium to agencies outside of Indiana. VACCINE has Memorandums of Understanding in place with all member agencies which permit the sharing of data, information, and materials needed to conduct research and development projects with the goals to coordinate and to improve public safety's analysis of data to promote improved operations in the law enforcement, fire, and emergency management and EMS fields. We are utilizing that data to create software and tools that aid the first responders in their duties.

We have deployed Purdue's Visual Analytics Law Enforcement Toolkit (VALET) to four law enforcement agencies so we can continue building improvements into the system that meet their needs. The Visual Law Enforcement Toolkit (VALET) provides law enforcement agencies with a suite of analytical tools coupled with an interactive visual interface for data exploration and analysis, forecasting, planning, and management. This system includes linked views and interactive displays that spatiotemporally model criminal, traffic and civil

(CTC) incidents and allows officials to observe patterns and quickly identify regions with higher probabilities of activity.

Our toolkit provides analysts with the ability to visualize different types of data sets (census data, daily weather reports, zoning tracts, prominent calendar dates, etc.) that provide an insight into correlations among CTC incidents and spatial demographics. In the spatial domain, we have implemented a kernel density estimation mapping technique that creates a color map of spatially distributed CTC events that allows analysts to quickly find and identify areas with unusually large activity levels. In the temporal domain, reports can be aggregated by day, week, month or year, allowing the analysts to visualize the CTC activities spatially over a period of time. Furthermore, we have incorporated temporal prediction algorithms to forecast future CTC incident levels within a 95% confidence interval. Such predictions aid law enforcement officials in understanding how hotspots may grow in the future in order to judiciously allocate resources and take preventive measures. Our system has been developed using actual law enforcement data and is currently being evaluated and refined by VACCINE's Public Safety Consortium of law enforcement agencies.

The VALET software is available on a PC, iPhone, or iPad and has been deployed to members of the VACCINE Public Safety Consortium including the Tippecanoe Co. Sheriff's Department, Lafayette Police Department, West Lafayette Police Department, and Purdue University Police Department. VALET provides advanced analysis capabilities, allowing analysts to develop and test hypothesis about criminal activities within various areas of their communities.



Valet on iPhone

We have deployed Georgia Tech's JigSaw, a relatively mature prototype system, within the West Lafayette and Lafayette Police Departments to begin looking at several unsolved cold cases. Until now, Jigsaw has provided more in the way of interactive visualization support of document exploration. In particular, Jigsaw visualizes connections between entities across documents to help investigators follow trails of information. Recently, we added enhanced computational analysis to the system. Jigsaw now also provides capabilities such as analysis of document similarity, document sentiment, document clusters by theme or content, and document summarization through a few words or sentences. Our focus has not been about developing innovative new algorithms for computational analysis, however. Instead, we have been exploring methods for smoothly integrating the computational analyses into an

interactive visual interface in a seamless manner that would provide a natural and fluid user experience.

We are working with the Lafayette and Indianapolis Fire Departments to develop a cell phone-based imaging system, MERGE (Mobile Emergency Response Guide), which can evaluate hazmat signage for the proper response protocols in the event of an accident involving hazardous materials.

Hazardous materials can react differently to stimuli and can cause problems in accidents and emergency situations. This fact makes them particularly dangerous to civilians and first responders. It is for this reason that the Emergency Response Guidebook was developed. This paper guide assists those in an emergency with the knowledge of how to handle Hazardous Materials. As one might expect, the guidebook is large and requires precious time to search an index to determine the best way to handle a particular hazardous material.

The MERGE system, shown below, is an electronic version of the guide with many new features and capabilities. These new capabilities include the use of image analysis methods to automatically determine the type of Hazardous Materials present based on an image taken of the sign/placard. MERGE has an easy to understand user interface to instruct an emergency responder or civilian in the proper way to handle a Hazardous Material emergency.



MERGE

MERGE was developed at the request of our first responders. The software is underdevelopment for use on iPhone and Android mobile devices. We anticipate that this should be deployed by summer 2011.  We have requests for use from the Lafayette, West Lafayette, Indianapolis, and St. Clair, MI Fire Departments.

We are also working with the Indianapolis Police Department's Gang Task Force on developing a mobile imaging application, GARI (Gang Graffiti Automatic Recognition and Interpretation)**,** that can interpret gang graffiti for police officers.

Gangs are a serious threat to public safety throughout the United States. Gang members are continuously migrating from urban cities to suburban areas. They are responsible for an increasing percentage of crime and violence in many communities. According to the National

Gang Threat Assessment, approximately 1 million gang members belonging to more than 20,000 gangs were criminally active within all 50 states and the District of Columbia as of September 2008. Criminal gangs commit as much as 80 percent of the crime in many communities according to law enforcement officials throughout the nation.

Street gang graffiti is their most common way to communicate messages, including challenges, warnings, or intimidation to rival gangs. It is, however, an excellent way to track gang affiliation and growth, or even sometimes to obtain membership information. The goal of this project is to use the knowledge gained from our work in mobile devices and applications and leverage it towards the development of a mobile-based system capable of image analysis. This system will provide an accurate and useful output to a user base through a database of gang graffiti images.

The image analysis includes obtainment of metadata (e.g., geoposition, date, and time) and the extraction of relevant features (e.g., color, shape) from the gang graffiti image. The information is sent to a server and compared against the graffiti image database. The matched results are sent back to the device where the user can then review the results and provide extra inputs to refine information. Once the graffiti is completely decoded and interpreted, it is labeled and added to the database.

GARI is under development for use on iPhone and Android mobile devices. We anticipate that this should be deployed by fall 2011. We have requests for use from the Lafayette, West Lafayette, Indianapolis, and Purdue Police Departments.



GARI

**Theme 2: Federal Operating Component Projects** (representative projects)
**United States Coast Guard**
VACCINE has spent the past year building strong ties within the USCG.  We have collaborated with them to build tools that can be adapted to real time situational areas.  Most significant of these projects was a detailed model allowing a thorough assessment of all CG search and rescue cases in the Great Lakes.  In the Coast Guard Search and Rescue Visual Analytics (CgSARVA) project, VACCINE presents a risk analysis system that enables the interactive visualization, analysis, and assessment of SAR missions completed by each USCG station.

The CgSARVA technology is deployed, and in use, at the USCG Atlantic Area Planning and Analysis Division. It will be deployed at the unit level during the summer of 2011.

The system enables the interactive analysis of trends, patterns, anomalies, and distribution of cases and associated sorties.  Additionally, this assessment tool enables the determination of potential increase or decrease in risk with a reallocation of a resource; as well as known increases or decreases in the response time.  The CgSARVA system features include:

- Linked calendar and line graph views for analyzing data patterns and distributions
- Interactive filter controls for assessing case distribution loads and locations
- Density estimation for hotspot generation
- Linked time slider controls for interactive temporal animation and exploration
- Interactive statistical summary tools for report generation

The CgSARVA technology is a vital component for analyzing risk assessment as Coast Guard staffing levels fluctuate due to budget changes, retiring volunteers, and a variety of other factors.  This system provides managers and analysts with a suite of tools for analyzing the distribution of previous search and rescue cases and a methodology for understanding the risk, efficiencies, and benefits involved with reallocation or reduction of resources.

In addition to the CgSARVA work, VACCINE has collaborated with the CREATE Center to provide an Analytical Visualization of the Boston PWCS Patrol Resource Allocation Environment, also known as PROTECT (Port Resilience Operational/Tactical Enforcement to Combat Terrorism).

This is a contract under our Basic Ordering Agreement. Using the previous two years of patrol routes, we are developing statistical models for loiter point visits.  Predictions of the loiter point visits will be computed and sent to the Coastguard to compare against actual visits during a time period in which the visits were unknown to the analysts.  The predictive efficacy of this model will then be addressed, and determination on the ability of hostiles to predict patrol routes based on current patterns will be assessed.  Sparse data will be handled by predicting the number of days between visits.  Similarly in utilizing the PROTECT model, two years worth of patrol routes will be generated from the PROTECT model.  Similar statistical models will be created and a predictive model will be adapted to the PROTECT routes.  Predictions of future loiter point regions will be generated from PROTECT and sent to us and our model will now assess the ability of hostiles to predict patrol routes based on potential PROTECT routes.

An interactive visual analytics environment will be created that will overlay patrol routes and potential RIN numbers for targets.  This environment will be developed for both desktop and hand-held platforms with a first generation tool looking at historical patrol route logs and critical infrastructure overlays.  The first generation tool will be deployed to a provisional set of members for analysis and feedback.

Another Coast Guard project under contract is an Analytical Visualization of the Port Arthur, TX Economic Impact Study (with CREATE).   This also is a contract under our Basic Ordering Agreement. In conjunction with the resources allocation tool, we will also look at creating economic impact models that will show the effect that potential catastrophes (natural

or manmade) will have. Our system will incorporate new models, showing the potential impact over the Port Arthur area utilizing census tract information, historical data, the output of the CREATE economic model and predictive values. Initial versions of the system will focus on specific threat scenarios as determined by the partnering institutions. Our analytic environment will allow views and analysis of economic impact over time, over space, and over geography to show linkages and causalities and enable a better understanding of the relationship of certain primary and secondary economic effects.

**United States Secret Service**
A research project in support of the U.S. Secret Service involves Disposable Cell Phone Analysis. Disposable cell phones have become popular in the marketplace. These devices are attractive to terrorist and criminal perpetrators because they allow a purchaser to have access to pre-paid cellular time without having to submit extensive personal information. The devices are also attractive because they are inexpensive and easy to use.

Disposable cell phones are intended for use over a limited period of time. Therefore, these devices frequently have their connectivity functionality disabled, not allowing data syncing with a personal computer. When the devices are used in the commission of criminal or terrorist activity, LE investigators are faced with a challenging forensic examination problem.

Current forensic examination methods on these devices are laborious and time consuming, the phones must be disassembled and analyzed or investigators must take photos of the screen to present evidence that is resident on the device. Neither of these methods allow an examiner to view all of the data, easily parse the data or examine files on the questioned device.

This effort is focusing on the demonstration and development of methods and tools that will allow an investigator to acquire call logs, contacts, pictures, videos and text messages from disposable cell phones.

**Federal Emergency Management Agency**
*Business Continuity Information Network*
Studies have shown that businesses risk failure if they are unable to reopen quickly after a disaster. To reduce these risks and improve communications within the business community, VACCINE partner, Florida International University, has developed the Business Continuity Information Network (pronounced "bee-kin"), a web-based service where local businesses, emergency management, and organizations that assist businesses can gather to share critical information and support continuity efforts before, during and after a disaster. The BCIN South Florida Release supports business recovery programs initiated by county participants including Broward, Miami-Dade, Palm Beach and Monroe County through their respective Emergency Management offices.

Available year-round as a public service, this business-to-business community network provides participating companies a tool to track their key employees and supply chain status, and locate needed recovery goods and services. The system facilitates professional organizations like Chambers of Commerce to assist their members, and helps government agencies to assess damage and prioritize recovery needs. BCIN reports, maps and shares

critical up to-the-minute information about infrastructure conditions and recovery efforts by working with County Emergency Management Offices and major private infrastructure providers. BCIN will localize and tailor relevant information to business managers who will use this information to better assess an event's impact to their facilities, employees, suppliers and customers.

BCIN's B2B network, which includes companies from different industry sectors, business and trade associations and NGOs, helps businesses locate disaster recovery resources, products, and services while allowing them to collaborate on recovery efforts. Businesses can use BCIN to report their operational status and assistance needs to county government and NGOs so these organizations can prioritize their relief efforts. Recovery dollars spent locally helps to keep local businesses open and mitigates employment loss. The BCIN South Florida Release supports business recovery programs initiated by county participants including Broward, Miami-Dade, Palm Beach and Monroe County through their respective Emergency Management offices. Private sector participants include Office Depot, Wal-Mart, the Greater Miami Chamber of Commerce and a variety of local businesses across industry sectors. BCIN is supported by the Department of Homeland Security, IBM and the National Science Foundation.

*A Data Integration Framework for Enhancing Emergency Response Situation Reports with Multi-Agency, Multi-Partner Multimedia Data*
Responders in the field (both public and private participants) are capturing footage, pictures and video, of the disaster area with mobile phones. This data is being collected but is not integrated in the incidence command systems where situation reports, incidence action plans, etc are being held. When these materials are entered into such systems, it's done so in a manual way which is time consuming and requires careful human review and management.

To support emergency management operations, Florida International University is using advanced data integration and visual analytics techniques to analyze situation reports, incident action plans, and other type of responder reports, pictures and video clips captured in the field and automatically augment/link the reports directly to relevant multimedia content obtain in the field.  Data integration techniques are being used to identify disaster specific keywords and will be used to identify the documents.

## Customs and Border Protection
*Query Performance of the Cargo Database (2010)*
Massive amounts of data about ocean borne cargo shipments are now gathered from disparate sources and organized into a unified analytic database. To be useful for daily operations, such a database needs to be able to rapidly respond to analysts' *ad hoc* queries. This scenario, interactive analytic queries in large datasets, is one that VACCINE has studied at Stanford, but we had not had the opportunity to analyze with data of the size and scope found in the cargo domain. Working with Barry Siegel and Greg Meckstroth at SSC-PAC San Diego, VACCINE conducted a study of an extract of unclassified data from the 5TB Cargo database. The Cargo database was interesting not only because of its size, but also because it collects data from over two dozen different sources.

The principal finding of the study is that while the current structure of the Cargo database is remarkably flexible in terms of its ability to absorb data from a wide array of sources, this

generality comes with a significant cost in terms of query performance. For the set of queries VACCINE was given as part of the study, Stanford worked out optimizations which improve query performance by factors of 2x-100x over the original Cargo database. This result is not surprising as database organization and query performance are interrelated. Good database design can only be done with intimate knowledge of the queries. Although security and privacy reasons precluded this study from having access to any actual queries, our findings still provide useful guidance for those analyzing real queries to restructure the Cargo database.

In summation, VACCINE's research, while focused on delivering technological value to our customers, is also able to extend the state of the art in many areas including the following:

- Theoretical work in the science of analytical reasoning
- Fundamentals of visual representations for multisource, uncertain, temporally evolving, and conflicting data
- Data representations, integration, quality assessment, cleaning, and reduction for massive datasets, streaming data, and multisource data integration
- Methods and novel visual analytic tools tailored to user, task, and device requirements
- Mobile visual analytics for real-time applications
- Novel spatiotemporal modeling, detection, and prediction techniques
- New multisource correlation and visual analysis techniques (e.g., linked animal and human health surveillance)
- Novel video analytics for tracking, event detection, and mobile video analytic applications
- Novel population distribution adaptive cluster detection and aberration detection techniques
- Novel synthetic data generation techniques for spatiotemporal data streams that match observed statistical properties of source data but are completely anonymized
- Adapted presentation methods for actionable decision making for a diverse set of users

## III. VACCINE Research Project Descriptions

These are classified according to PI school location

*The rest of this page is intentionally left blank*

**School**

Florida International University

a) **Project Name /Theme**

A Data Integration Framework for Enhancing Emergency Response Situation Reports with Multi-agency, Multi-Partner Multimedia Data;

Theme 1: Public Safety Coalition Projects

b) **PI/CO-PI and Institution**

Dr. Shu-Ching Chen
PI, Professor, School of Computing and Information Sciences, Florida International University

Dr. Tao Li
Co-PI, Associate Professor, School of Computing and Information Sciences, Florida International University

Steven Luis
Senior Investigator, School of Computing and Information Sciences, Florida International University

**Collaborating End-Users**

| Paul Vitro Emergency Management Coordinator Miami-Dade County Department of Emergency Management 9300 NW 41St. Miami, FL 33178 Phone: 305-468-5423 www.miamidade.gov/oem | Ray Misomali, Emergency Management Coordinator Miami-Dade County Department of Emergency Management 305-468-5429 www.miamidade.gov/oem | Soheila Ajabshir - GISP, FPEM Systems Manager Miami-Dade County Department of Emergency Management (DEM) 9300 NW 41 St., Doral, FL 33178 (305) 468 - 5417 |
|---|---|---|

c) **Research Problem (Abstract)**

In this project, we propose to use advanced data integration and visual analytics techniques to systematically augment situation reports which include pictures and text with relevant multimedia content obtained in the field. Our proposed framework implements the offline Multimedia Analysis Framework (MAF) to associate situation reports with relevant multimedia content. The MAF framework includes a Hierarchical image classification component developed based on multi-source data fusion and some data mining techniques, such as multiple correspondence analyses (MCA). The situation report is pre-processed by using the GATE system and WordNet to associate location-subject pairs. We also provide the prototype with applications to emergency response. Specifically, an iPad application based on the proposed

framework has been developed using the pre-defined ontology with the incorporation of domain knowledge. Furthermore, a system improvement mechanism based on user feedback is designed to refine the association between situation report and multimedia data. Finally, the proposed system was evaluated by collaborating with EOC department of Miami-Dade, where the personnel have expressed interest in the application and encouraged us to further develop the system into an operational pilot.

d) **Technical Approach**

Data Collection—Data Mining, Expert consultation; Analytic Methods—Case Studies, Modeling; Nature of Research—Applied, Coordination/Integration

e) **Deliverables/Milestones**

Multimedia Analysis Framework (MAF);

Novel techniques for integrating clustering and dimensionality reduction for visualization;

Outlier removal and principal component subspace projection algorithms;

Hierarchical image classification component based on multi-source data fusion and Multiple Correspondence Analysis (MCA);

Prototype system with applications to emergency response;

Techniques for incorporating domain knowledge and user interaction;

Document summarization and visualization methods.

System evaluation form

f) **Project Outcomes (Results, Accomplishments, Conclusions)**

**Publications**

Lei Li, Wenting Lu, Jingxuan Li, Tao Li, Honggang Zhang and Jun Guo, "Exploring Interaction Between Images and Texts for Web Image Categorization," In Proceedings of The 24th Florida Artificial Intelligence Research Society Conference (FLAIRS-24), 2011.

Yimin Yang, Hsin-Yu Ha, Fausto Fleites, Shu-Ching Chen, and Steven Luis, "Hierarchical Disaster Image Classification for Situation Report Enhancement," accepted by The 12th IEEE International Conference on Information Reuse and Integration (IRI 2011), Las Vegas, Nevada, USA, August 3-5, 2011.

**Presentations**

2/22/2010, Jason Allen, "Enhancing Situation Documents with Disaster Related Images," Overview of Prototype System at Miami-Dade County Department of Emergency Management.

3/9/2011, Jason Allen, "Enhancing Situation Documents with Disaster Related Images," Public Safety Consortium Meeting at Purdue University.

4/11/2011, Jason Allen, "Enhancing Situation Documents with Disaster Related Images," Evaluation Activity at Miami-Dade County Department of Emergency Management.

**Patents/Copyrights**
System software is copyrighted Florida International University.

**Research Assistant**
Li Zheng, School of Computing and Information Sciences, Florida International University

**Other Outcomes/impacts**
A preprocessing Multimedia Analysis Framework (MAF) has been developed for associating situation reports with relevant multimedia content, in which a Hierarchical image classification component was designed and developed based on multi-source data fusion and multiple correspondence analyses (MCA) techniques. The document pre-processing is completed by using the GATE system and WordNet to associate location-subject pairs with situation reports. A system improvement mechanism based on user feedback has been designed to refine the association between situation reports and multimedia data. Finally, an iPad application based on the proposed system has been developed and presented to the personnel at Miami-Dade Emergency Management (MDEM), where an evaluation activity regarding the developed application was successfully conducted and the people there have expressed interest for us to further develop the system into an operational pilot. In the future, more extensive disaster ontology for categorizing images will be designed and developed; images and situation reports will be associated in real-time for online analytic purposes; Improvements will be made to highlight the discovered entities so that it is easier for a user to understand the relationship between the document and images selected. A more intelligent and comprehensive user interface will be designed for a better user experience.

**School**
Georgia Institute of Technology

a) **Project Name/Theme**
Visual Analytics for Investigative Analysis on Text Documents

Theme 1: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**
John Stasko, Georgia Institute of Technology (PI)

**Collaborating Partners**
National Science Foundation (FODAVA)

**Collaborating End-Users**
Cindy Marion, West Lafayette Police Dept., (765) 775-5200,
Steve Hawthorne, Lafayette Police Dept., (765) 807-1227, sdhawthorne@lafayette.in.gov
Richard LeCates, Clayton County Police Dept, (770) 473.5848, Rich.LeCates@co.clayton.ga.us

c) **Research Problem (Abstract)**
We are developing techniques and systems for investigative analysis on large collections of documents including both unstructured text and structured documents (e.g., spreadsheets). More specifically, we have developed the Jigsaw visual analytics system to help investigators perform exploration, sense-making, and understanding tasks on documents. By helping analysts to determine which documents to read first and to find important connected documents, we can help them to make more timely and accurate assessments of the situations described within those documents. The project is continuing the development of Jigsaw and exploring how new capabilities can be added to it. We also disseminate the system for use by people and organizations in various law enforcement, intelligence, and security domains.

d) **Technical Approach**
Analytic Methods—Case Studies, Sampling; Nature of Research—Applied, Basic, Coordination/Integration, Hybrid Basic-Applied

e) **Deliverables**

| Milestones for VACCINE Year 2 | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| Submit paper to HCIR Workshop. | X | X | | |
| Implement computational analysis techniques in Jigsaw. | X | X | | |
| Create trial distribution of system with new analyses. | | X | | |
| Visit potential users in Washington, D.C. | | X | | |
| Submit paper to CHI | | X | | |
| Demonstrate system at BC (Canada) Justice Inst. Meeting. | | X | | |
| Create new official distribution of system for external use. | | | X | |

| | | | | |
|---|---|---|---|---|
| Create tutorial videos about Jigsaw. | | | | X |
| Demonstrate new version of system at VACCINE Annual meeting | | | X | |
| Continue distribution and working with external clients. | X | X | X | X |
| Create new database import capabilities. | | | | X |
| Travel to Indianapolis to learn about police work there | | | X | |
| Travel to Indianapolis to provide tutorial on system use and gather feedback | | | X | |
| Provide Jigsaw to local police in Indianapolis | | | X | |

f) **Project Outcomes (Results, Accomplishments, Conclusions)**

**Publications**

Carsten Görg, Jaeyeon Kihm, Jaegul Choo, Zhicheng Liu, Sivasailam Muthiah, Haesun Park, and John Stasko, "Combining Computational Analyses and Interactive Visualization to Enhance Information Retrieval", 2010 Workshop on Human-Computer Interaction and Information Retrieval, New Brunswick, NJ, August 2010.

C. Görg, H. Tipney, K. Verspoor, W.A. Baumgartner Jr, K.B. Cohen, J. Stasko, L.E. Hunter, "Visualization and Language Processing for Supporting Analysis Across the Biomedical Literature", Proceedings of International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES), September 2010, pp 420-429.

Zhicheng Liu, Carsten Görg, Jaeyeon Kihm, Hanseung Lee, Jaegul Choo, Haesun Park, John Stasko, "Data Ingestion and Evidence Marshalling in Jigsaw VAST 2010 Mini Challenge 1 Award: Good Support for Data Ingest", Proceedings of IEEE VAST '10, Salt Lake City, UT, October 2010, pp. 271-272.

Hanseung Lee, Jaegul Choo, Carsten Görg, Jaeeun Shim, Jaeyeon Kihm, Zhicheng Liu, Haesun Park, John Stasko, "GeneTracer: Gene Sequence Analysis of Disease Mutations VAST 2010 Mini Challenge 3 Award: Excellent Process Explanation", Proceedings of IEEE VAST '10, Salt Lake City, UT, October 2010, pp. 291-292.

Zhicheng Liu and John T. Stasko, "Mental Models, Visual Reasoning and Interaction in Information Visualization: A Top-down Perspective", *IEEE Transactions on Visualization and Computer Graphics,* (Paper presented at InfoVis '10), Vol. 16, No. 6, November/December 2010, pp. 999-1008.

Youn-ah Kang, Carsten Görg, John Stasko, "How Can Visual Analytics Assist Investigative Analysis? Design Implications from an Evaluation," *IEEE Transactions on Visualization and Computer Graphics,* Vol. 17, No. 5, May 2011, pp. 570-583.

**Presentations**

NSF III Program Annual Review Meeting, Arlington, VA, April 2010 "Visual Analytics", John Stasko

AAAS Symposium on Novel Methods for the Evaluation of Federal Research Programs, Washington, D.C., April 2010

"Visual Analytics for Exploring and Investigating Research Programs", John Stasko NSF, Arlington, VA, August 2010

"Jigsaw – Visual Analytics for Data & Document Exploration and Understanding", John Stasko

Justice Institute of British Columbia, Vancouver, B.C., September 2010 "Investigative Analysis of Document Collections with Jigsaw", John Stasko

SUNY Stony Brook, Computer Science Distinguished Lecture Series, February 2011

"Visual Analytics for Investigative Analysis and Exploration of Documents and Data", John Stasko

Investigative Reporters and Editors, Computer-Assisted Reporting Conference, Raleigh, NC, February 2011

"Jigsaw", John Stasko CNN, Atlanta, GA, March 2011 "Investigative Analysis of Document"

**Scholarly Journal Citations**
Six

**Research Assistants**
Zhicheng Liu, Georgia Tech, School of Interactive Computing
Youn-ah Kang, Georgia Tech, School of Interactive Computing

**Follow-on Funding/Related Projects**
National Science Foundation, FODAVA Program, PI: Haesun Park (Stasko is a co-PI)
8/08-7/13

**Other Outcomes/Impacts**
Our progress developing and distributing the Jigsaw system has been very good. With respect to the actual system development, we have added new computational analysis capabilities to the system, new visualizations, and new data import capabilities. We published papers at the HCIR Workshop and the KES Conference about our research. Additionally, we submitted and had accepted a journal paper to *IEEE Transactions on Visualization and Computer Graphics* about a user study on Jigsaw. We used Jigsaw to work on the IEEE VAST Conference Challenge and won an award for our efforts.

We created a new release (0.35) of the system and began distributing it to people and organizations who want to use it. To accompany the system, we have created new tutorial videos that illustrate usage of the different components of the system, as well as improved written tutorial documentation. We have a growing and wide-ranging client set now and we continue to work on expanding this. We are pursuing collaboration with local police departments in Georgia and Indiana specifically. We are also exploring new document domains for analysis such as bioinformatics, business markets, on-line reviews, and academic papers.

Doctoral student Youn-ah Kang studied teams of student analysts at Mercyhurst University in Erie, PA in order to better understand their tasks and requirements. Her analysis should assist the developers of visual analytics systems for the intelligence community. Doctoral student Zhicheng Liu has begun the system of a system called Ploceus for analysis of tabular, relational data through networked visualizations.

John Stasko gave a number of talks over the year about visual analytics in general and the Jigsaw system more specifically. Talk locations included the National Science Foundation, the Justice Institute of British Columbia, the Computer-Assisted Reporting Conference of the Investigative Reporters and Editors Association, CNN, and an AAAS Symposium.

**School**
Pennsylvania State University

a) **Project Name/Theme**
Symbology Standardization Support

Theme 2: Federal Operating Component Projects

b) **PI/Co-PI and Institution**
Alan MacEachren, Pennsylvania State University (PI)
CoPI: Anthony Robinson, Pennsylvania State University

**Collaborating End Users**
CBP (Customs and Border Patrol)
FEMA (Federal Emergency Management Agency)
IICD (Infrastructure Information Collection Division)
GMO (Geospatial Management Office)
DHS internal Geospatial Working Group

c) **Research Problem (Abstract)**
The Department of Homeland Security contains many diverse organizations that produce or use maps in daily operation. Audiences for these maps range from geospatial analysts to the general public. While individual organizations have their own standards or conventions for map symbology, there is not a consistent set of symbols used across DHS. To support information sharing across the Department, DHS S&T sponsors this project to document the use of map symbols in homeland security missions and to develop a repeatable process for expanding symbol sets to support new needs. In addition, the project is developing a web-based *Symbol Store* that will permit users to request map symbols to support their specific needs, enable DHS agencies to contribute symbol sets to share, and to act as the hub for efforts to eliminate inconsistencies in map symbology within and across agencies and to improve map symbology across DHS so that maps provide maximum support to DHS missions.

To date, the project team has worked with personnel from Customs and Border Patrol (CBP), Infrastructure Information Collection Division (IICD), United States Coast Guard (USCG), United States Fire Service (USFS), National Operations Center (NOC), Federal Emergency Management Agency (FEMA), and the Domestic Nuclear Detection Office (DNDO) to assess map symbology use (during 2009-2010); develop and carry out a Symbology Standardization Process with personnel from three agencies (as separate processes): CPB (during 2009-2010), FEMA and IICD (during 2010-2011); study issues related to dynamic and multiscale map symbols, and develop and implement the initial prototype Symbol Store to distribute approved map symbols among DHS personnel.

d) **Technical Approach**

| Data Collection Methods | Analytic Methods | Nature of Research |
|---|---|---|
| Compiling & Sorting DB<br>Expert consultation<br>Survey<br>Other: Structured interviews w/practitioners | Case Studies<br>Sampling<br>Statistical analysis | Hybrid Basic-Applied<br>Hybrid Applied –Consultation<br>Coordination/integration |

e) **Deliverables**

Symbology project poster:
http://www.geovista.psu.edu/resources/flyers/MacEachrenEtAl_2010_TCIP.pdf
Symbology project poster #2:
http://www.geovista.psu.edu/symbology/Symbology_Overview_Poster.pdf
Symbol Store prototype video: http://www.youtube.com/watch?v=U0OA0X5m4qc
e-Symbology Portal: a customized Drupal content management system
Symbol Store working prototype (currently available to DHS project management team and to DHS GMO)

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Published during project year

Robinson, A.C., Roth, R. E., and A.M. MacEachren (2010) Challenges for Map Symbol Standardization in Crisis Management. *7th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. Seattle, WA, May 2-5, 2010.

Roth, R.E., Finch, B.G., Blanford, J.I., Klippel, A., Robinson, A.C. and MacEachren, A.M. 2010: The card sorting method for map symbol design. *AutoCarto 2010*, Orlando, FL: ASPRS, Nov. 15-19, 2010.

**Accepted**
Robinson, A.C., Roth, R.E., and A.M. MacEachren (in press) Understanding User Needs for Map Symbol Standards in Emergency Management. *Journal of Homeland Security and Emergency Management 4.*

Roth, R.E., Finch, B.G., Blanford, J.I., Klippel, A. Robinson, A.C. and MacEachren, A.M. (in press) Card sorting for cartographic research and practice. *Cartography and Geographic Information Science* 38(2).

**Presentations**
Robinson, A.C., Pezanowski, S., Hardisty, F., Roth, R.E., Blanford, J., and A.M. MacEachren. "A Web-Based Symbol Store for Sharing Map Symbology" Presentation at the annual meeting of the North American Cartographic Information Society in St. Petersburg, FL. October 13-15, 2010.

Robinson, A.C. "Challenges to Standardizing Map Symbology" Presentation at the annual meeting of the Association of American Geographers in Washington, DC. April 14-18, 2010.

**Presentations to be given before 5/31/2011:**
Bianchetti, R.A., J. Blanford , A. Klippel, A.C. Robinson, and A.M. MacEachren. 2011. Critical Evaluation of International Emergency Management Symbology. The 107th Annual Meeting of the Association of American Geographers. Seattle, WA: April 15, 2011.

Guidero E, F Hardisty, RE Roth, A Klippel, AC Robinson, and AM MacEachren. 2011. Representing change using point symbols. The 107th Annual Meeting of the Association of American Geographers. Seattle, WA: April 15, 2011.

Pezanowski S, AC Robinson, E Guidero, R Bianchetti, F Hardisty, J Blanford, RE Roth, and AM MacEachren. 2011. An interactive, web-based tool for discovering and sharing map symbols. The 107th Annual Meeting of the Association of American Geographers. Seattle, WA: April 15, 2011.

**Reports**
One thus far; initial publications have just started to appear

**Research Assistants**
Robert Roth, Penn State, Geography (PhD candidate)
Raechel Bianchetti, Penn State, Geography (PhD candidate)
Elaine Guidero, Penn State, Geography (MS candidate)

Benjamin Finch, Penn State, Geography; graduated fall 2010 (undergraduate)

**Technology Transitions**
For project year 2009-2010 we provided DHS with a full codebase for our customized e-Symbology web portal (based on Drupal). We also included documentation to help DHS launch their own version, and results from the first round of e-Symbology testing from collaboration with mapmakers at Customs and Border Patrol.

One of three tasks within the work detailed for 2010-2011 is to develop a prototype of the Symbol Store as a web-based application. We have developed a working version and done multiple briefings for personnel in DHS-GMO, the DHS Geospatial Working Group, and contractors who will be responsible for transitioning the Symbol Store into practice within DHS. Plans for supporting that transition process while continuing development of Symbol Store capabilities and integrating aspects of the Symbology Standardization Process within it have been developed for work during a funding period proposed to begin July 1, 2011 and continue through March 31, 2013.

**Other Outcomes/Impacts**
*Introduction*
Our symbology research for DHS to date has resulted in several key research contributions. We have developed a new symbol standardization process based on needs identified in a series of interviews with DHS mapmakers. To support this process, we designed and implemented the *e-Symbology Portal*, which is a customized Drupal content management system that allows DHS

users to participate in our iterative symbol standardization process at a distance, and asynchronously. We have tested this process so far with three DHS components (CBP, FEMA, and IICD). Lessons learned from our efforts to test our process with DHS mapmakers led us to develop the Symbol Store prototype, which allows users to browse, search for, and download map symbols through an intuitive web-based interface. Finally, we have made progress on initial efforts to explore and characterize the design constraints and opportunities associated with Multi-Scale and Dynamic symbols (report to be submitted in June, 2011). The following sections provide more details on each contribution, and we conclude with a brief section outlining future research directions.

*Symbol Standardization Process*
Based on interviews with DHS mapmakers conducted in the first year of our project, it became clear that a new symbol standardization process to formalize, refine, and share mission specific map symbol standards would be the best direction for our project. Such an approach would allow DHS units to create maximally-useful symbol standards and set the stage for comparing the full range of mission specific standards to identify key commonalities and distinctions in future DHS symbol interoperability efforts.

The standardization process we developed relies on distributed, asynchronous methods so that busy professionals are not required to gather in person at the same time (a key constraint when working with busy DHS users). Our approach makes use of flexible open-source web tools to support and capture the process of standard development. This strategy enables repeatability and helps to document key decision points and the rationale for them to reveal further insights into the use and development of map symbol standards.

*e-Symbology Portal*
Our process development platform, called the e-Symbology Portal, is based on a customized Drupal (http://drupal.org/) content management site in which we have created round-based activities to iteratively refine, augment, and formalize a current mission-area map symbol standard. Participants who use our platform can contribute in each round within a specified timeframe (1-5 days, depending on the activity). Contributions in each round are moderated to distill feedback into key issues for further discussion and voting. In addition, we have implemented a procedure to anonymize participation to promote diversity of opinions - similar to the way in which a Delphi exercise functions.

To evaluate and refine our process for symbol standard development, we worked (during 2009-2010) with participants at Customs and Border Patrol (CBP) to formalize their current "in-house" map symbol standard. The results of this study (and two subsequent studies this year with participants from FEMA and IICD) revealed a large number of symbol design, duplication, and category issues that can be rectified. Participants expressed satisfaction with both the process itself as well as the e-Symbology platform that they used to complete the process.

*Symbol Store*
Our work to develop and evaluate a symbol standardization process revealed the need for a simple technical mechanism for sharing symbols among mapmakers at DHS. To meet DHS needs, we have designed and implemented a prototype interactive web interface for contributing, browsing, and sharing symbology that we call the Symbol Store (figure 1). While a single common symbology for all DHS missions may not be realistic, it is quite worthwhile to collect

and make available all mission specific symbology in a single place, where symbols can be tagged with topic and agency keywords to support natural language search. For example, a mapmaker at FEMA can go to the Symbol Store site to look up all the symbols currently in use across DHS to represent Fire Stations. The Symbol Store results page can display all of the relevant symbols, including the range of keywords that describe them as well as tags identifying which DHS mission areas are using which symbols. The FEMA mapmaker then can interactively select which symbols they want to download and the Symbol Store will offer the selected set in a GIS-compatible format.

In addition to supporting interoperability among mapmakers at DHS, the Symbol Store affords unique opportunities for us to track and evaluate the most frequently used symbols, to identify duplicate symbols, and to explore how symbol categories differ among different DHS mission areas.

*Multi-scale and Dynamic Symbology*
Our focus on Multi-scale symbols has been toward the goal of developing a methodology and technique for developing and organizing scale-dependent map symbology and rules for generalization between scales. We have completed a content analysis of approximately one hundred DHS maps to characterize current methods for implementing cartographic symbol generalization at different scales. This content analysis study has identified map features, scales, and thematic content and summarized commonalities between the maps. The results of this work will be used to create base maps for the Symbol Store and to present a set of common mapping practices across multiple DHS organizations.

We have also been working on determining ways to enable dynamic symbology for mapmakers at DHS. We have begun this task by using cartographic and design theory to determine which dynamic symbolization strategies are most effective. This has taken the form of a user study, for which we have developed custom software to conduct the study with student participants who will use this software to rate different symbol designs and describe how well different dynamic symbol design strategies enable (or hinder) interpretation. The end result of this work will be empirical evidence to help guide design decisions for creating new dynamic symbols (for example, should we use color to indicate changes to damage levels, or would texture be more interpretable in most circumstances?).
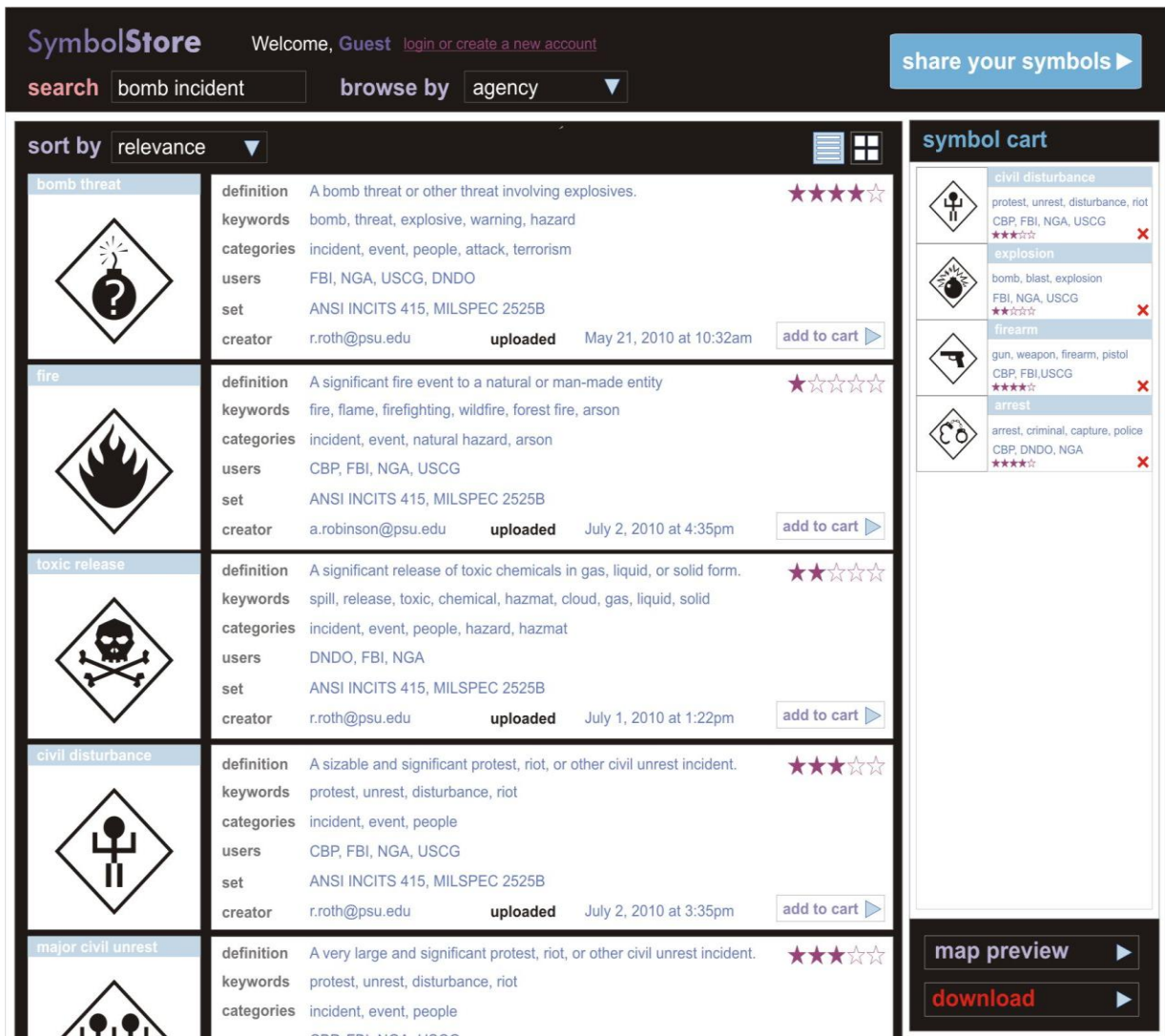


**Figure 1. Symbol Store prototype interface. Users can search by concept to find potentially relevant symbols, then drag those that are suitable to the "symbol cart" for preview on a map (functionality still in development) or download to an ESRI sylefile.**

*Future Directions*
Our next steps include further development of the Symbol Store, emphasizing its usability and utility for real world end-users at DHS. We will be conducting multiple evaluations to refine the Symbol Store interface and to test it in situations that mirror real-world use contexts. In addition,

we will be focusing on creative ways to enable users to discuss and vote on symbols to be included in future DHS symbol standards through the Symbol Store interface itself (thus bringing together what we have built in the e-Symbology Portal with the Symbol Store prototype). Further, we will explore the semantic issues of concepts that may be similar or identical but identified by different labels in different agencies and similar issue of concept/feature category systems that do not match and the implications of this for map symbology that is usable across agencies in mission situations (e.g., managing a crisis) in which multiple agencies need to share geographic information and develop integrated plans.

We will also continue to identify and evaluate best practices for designing multi-scale and dynamic symbology, with a special emphasis on ensuring that symbols will function appropriately in mobile GIS environments on smartphones and tablets.

Finally, we will be looking for opportunities to bring what we have learned so far from symbology research with DHS to other related government and non-government stakeholders that have similar emergency management mapping needs.

**School**
Pennsylvania State University

a) **Project Name/Theme**
MDRP 3 – Situational Surveillance & In-Field Criminal Investigative Analytics: Geovisual Analytics Methods

SP 14-Task 3: Technology Integration into ExPod at JIBC

MDRP 13: GeoJunction: Collaborative Visual-Computational Information Foraging and Contextualization to Support Situation Awareness

Theme: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**
Alan MacEachren, Pennsylvania State University (PI)

**Collaborative Partners**
PSU CoPI: Prasenjit Mitra, Pennsylvania State University
Purdue Collaborators: David Ebert & Tim Collins (for VACCINE Public Safety Consortium)
Related support for follow-on research provided as option-year funding on a grant from the National Geospatial-Intelligence Agency
Related support for follow-on research provided through a National Science Foundation Contract
Related support for follow-on research pending for a U.S. Army Corps of Engineers contract

**Collaborative End-Users**
Harrisburg Bureau of Police – Contact: Captain Cindy Baldwin: CBaldwin@cityofhbg.com
VACCINE Public Safety Consortium – Contact: Tim Collins: tfcollins@purdue.edu

c) **Research Problem (Abstract)**
*MDRP 3 & SP 14:*
This research and associated transition activity is focused on developing geovisual analytics methods and tools that support situational surveillance and in-field criminal investigative analytics. The emphasis is on enhancing the ability of analysts to leverage location and time components of data about crime (and related data relevant to solving and preventing crime). The research has followed an iterative, user-centered design process. The process began with in-depth interviews with law enforcement personnel at local, state, and national level. The insights gained were integrated into CrimeViz, a client-server based web-mapping/analysis application that is being iteratively refined with regular input from our partner agency, the Harrisburg, PA Bureau of Police. CrimeViz is being transitioned for their use and the software has been shared with collaborators at Purdue who plan to add capabilities and help facilitate wider distribution through the Public Safety Consortium. SP-14 work has been focused specifically on assessing the potential to develop a training scenario using CrimeViz for use in the ExPod at JIBC (initial reactions have been positive).

*MDRP 13 & SP 15:*
This project focuses on developing, implementing, assessing, and transitioning methods and tools to (a) help analysts, crisis managers, and policy makers find and contextualize relevant

heterogeneous information quickly to (b) support the situation awareness (SA) required to interpret evidence and make decisions in a continually changing environment. Specifically, the project is focused on developing knowledge-enabled, place-time aware, and visually-computationally enhanced web-based applications that support: (a) open source information foraging, entity and relation extraction, and contextualization; (b) document queries filtered by place, time, person/organization, and concept; (c) exploration of connections between social and geographic networks; (d) multi-scale analysis; and (e) situation assessment and monitoring from situation reports, news stories, social media and other sources.

The research has produced multiple web-based applications addressing components of the challenges. These include: Health GeoJunction (supporting place, time, concept filtering of scientific documents relevant to health crises), the Context Discovery Application (supporting construction of geo-historical context for crisis relief decision making), SensePlace (supporting information foraging in open source documents), and SensePlace2 (supporting use of social media micoblogs as input to situation awareness for crisis management); sork on SensePlace2 was the primary focus of SP-15. Funding to extend applications has been provided by NGA and NSF; and a Gates Foundation-funded project provided a test case of SensePlace for information foraging as input to infectious disease modeling.

d) **Technical Approach**

| Data Collection Methods | Analytic Methods | Nature of Research |
|---|---|---|
| Compiling & Sorting DB<br>Data Mining<br>Expert consultation<br>Survey | Case Studies<br>Sampling<br>Statistical analysis<br>Modeling<br>Risk assessment<br>Sampling | Consultation<br>Hybrid Basic-Applied<br>Hybrid Applied –Consultation<br>Coordination/integration |

e) **Deliverables**
**Videos Produced to Demonstrate Software**
Brian Tomaszewski, Alan M. MacEachren, 2010. Geo-Historical Context Support for Information Foraging and Sensemaking Implemented in the Context Discovery Application,

**Video Posted on Youtube**, April 16, 2010 (video production assistance provided by Krista Kahler; produced to accompany Tomaszewski, B. and MacEachren, A.M. 2010: Geo-Historical Context Support for Information Foraging and Sensemaking: Conceptual Model, Implementation, and Assessment. *IEEE Conference on Visual Analytics Science and Technology (IEEE VAST 2010),* Salt Lake City, Utah, USA, 139-146. http://www.youtube.com/watch?v=TGCw9YlPqLc

Alan M. MacEachren, Scott Pezanowski, 2010. Health GeoJunction, **Video Posted on Youtube**, April 30, 2010 (video production assistance provided by Krista Kahler, produced to accompany A. M. MacEachren, M. Stryker, I.J. Turton, S. Pezanowski, "HEALTH GeoJunction: Visual-computationally Enabled Browsing of Health Publications and Their Contents," International Journal of Health Geographics, vol. 9, no. 23, 2010.) http://www.youtube.com/watch?v=pauyUrAn7lg

Robert Roth, Alan M. MacEachren, 2010. GeoVISTA CrimeViz Update, **Video Posted on Youtube**, Oct 15, 2010. http://www.youtube.com/watch?v=Wj-aBS1hQ7g

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Published during project year

Andris, C. and Hardisty, F. (2011). Extended Abstract: *Visualizing Migration Dynamics Using Weighted Radial Variation*. GeoViz Hamburg. Hamburg, Germany, March 10-11, 2011.

Auer, T., MacEachren, A.M., McCabe, C., Pezanowski, S. and Stryker, M. 2011: HerbariaViz: A web-based client-server interface for mapping and exploring flora observation data. *Ecological Informatics*, doi:10.1016/j.ecoinf.2010.1009.1001.

Hardisty, F. and Klippel, A. 2010: Analyzing spatio-temporal autocorrelation with LISTA-Viz. *International Journal of Geographical Information Science* 24, 1515 - 1526.

Hardisty, F. and Robinson, A.C. 2010: The Geo Viz toolkit: using component-oriented coordination methods to aid geovisualization application construction. *International Journal of Geographical Information Science*, DOI: 10.1080/13658810903214203

MacEachren, A.M., Stryker, M.S., Turton, I.J. and Pezanowski, S. 2010: HEALTH GeoJunction: Visual-computationally Enabled Browsing of Health Publications and Their Contents. *International Journal of Health Geographics* 9.

Griffin, A.L. and Robinson, A.C. 2010: Comparing Color and Leader Line Approaches for Highlighting in Geovisualization. In Purves, R. and Weibel, R., editors, *Extended Abstracts Volume, GIScience 2010,* Zurich, Switzerland: GIScience, 14-17th September, 2010.

Robinson, A. 2010: Supporting Synthesis in Geovisualization. *International Journal of Geographical Information Science*.

Roth, R.E., Ross, K.S., Finch, B.G., Luo, W. and MacEachren, A.M. 2010: A user-centered approach for designing and developing spatiotemporal crime analysis tools In Purves, R. and Weibel, R., editors, *Extended Abstracts Volume, GIScience 2010,* Zurich, Switzerland: GIScience, 14-17th September, 2010.

Tomaszewski, B. and MacEachren, A.M. 2010: Geo-Historical Context Support for Information Foraging and Sensemaking: Conceptual Model, Implementation, and Assessment. *IEEE Conference on Visual Analytics Science and Technology (IEEE VAST 2010),* Salt Lake City, Utah, USA, 139-146, October 24 - 29, 2010.

Tomaszewski, B., J. Blanford, K. Ross, S. Pezanowski and A. MacEachren. 2011. "Supporting Rapid Sensemaking in Diverse Web Document Foraging." *Computers, Environment and Urban Systems* 35: 192-207.

Tomaszewski, B. 2011: Situation awareness and virtual globes: Applications for disaster management. *Computers & Geosciences* 37, 86-92.

White J.D.D. and R.E. Roth. 2010. TwitterHitter: Geovisual analytics for harvesting insight from volunteered geographic information. In: *Proceedings of GIScience 2010*. Zurich, Switzerland, 14-17th September, 2010.

**Accepted**
MacEachren, A.M., Robinson, A.C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J. and Mitra, P. in press: Geo-Twitter Analytics: Applications in Crisis Management. *25th International Cartographic Conference,* Paris, France, July 3-8, 2011.

Caragea, C., McNeese, N., Jaiswal, A., Traylor, G., Kim, H.W., Mitra, P., Wu, D., Tapia, A.H., Giles, L. and Jansen, B.J. 2011: Classifying Text Messages for the Haiti Earthquake. *Proceedings of the 8th International Information Systems for Crisis Response and Management (ISCRAM) Conference*, Lisbon, Portugal. K. Bajpai, A. Jaiswal, "A Framework for Analyzing Collective Action Events on Twitter", *8th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, 2011

**Presentations**
Hardisty, F. 2010a. Interactive Syndromic Surveillance of Influenza Rates using the Geo Viz Toolkit. Analysis, Visualization and Reporting (AVR) Webinar CDC.

Hardisty, F. 2010b. Open Source Geovisual Analytics. GIS-Pro. Orlando, FL, URISA.

Hardisty, F. 2010c. Using text analytics with spatial statistics to characterize H1N1 flu outbreaks. *The 106th Annual Meeting of the Association of American Geographers.* Washington, DC.

Hardisty, F. and Klippel, A. (2010). Analyzing Spatio-Temporal Autocorrelation with LISTA-Viz. GeoVA(t) – Geospatial Visual Analytics: Focus on Time. Guimarães, Portugal, ICA Commission on GeoVisualization.

Hardisty, F. (2011). Exploring Geospatial Data with the GeoVizToolkit. North Carolina GIS User's Group Meeting.

Hardisty, F. and Kroon, J. (2011). Cloud-Enabled Space-Time Clustering. AAG Annual Meeting. Seattle.

Hardisty, F., Peuquet, D., Xu, S. and Robinson, A.C. (2011). Methods for Ad-hoc Delineation and Analysis of Categories of Spatio-Temporal Events. COM.Geo. Washington, D.C., ACM Digital Library.

MacEachren, A.M. Geovisual Analytics: Supporting visual reasoning with complex information, Annual Meeting of the AAG, April 14-18, 2010 – Washington, DC.
MacEachren, A.M. 2010. Geovisual Analytics for Crisis Management: Moving Beyond GIS. Invited keynote at Information Systems for Crisis Response and Management (ISCRAM) 2010, May 2-5, 2010, Seattle, WA.

MacEachren, A.M. 2010. Geovisual Analytics for Health, panel presentation at the Visual Analytics Community (VAC) Consortium Conference, Aug. 31 – Sept. 1, 2010, University of Maryland, College Park MD.

Murdock, M.J., Maziekas N.V., and R.E. Roth. 2011. Geovisual analytics of improvised explosive device incidents. *The 107th Annual Meeting of the Association of the Association of American Geographers*. Seattle, WA: April 12-16.

Mitra, P., "Information Extraction for Geospatial Visual Analytics." University of Houston Downtown, Computer Science Seminar, Houston, 2010.

Murdock M, B.G. Finch, N. Maziekas, and R.E. Roth 2010. The basic ordinance observational management system: Geovisual analytics of improvised explosive device incidents. *The Association of American Geographers, Middle States Division 2010 Meeting.* West Point, NY: October 22-23.

Roth R.E., KS Ross, W Luo, BJ Finch, and AM MacEachren. 2010. A user-centered approach to extending spatiotemporal crime analysis tools. *The 106th Annual Meeting of the Association of American Geographers*. Washington, DC: April 15.
Roth R.E. 2010. Geovisual analytics for crime analysis and homeland security *VACCINE Consortium Technology Demo Day*. West Lafayette, IN: July 14.

Roth R.E. 2011. Useful and usable geovisualization for crime analysis. *The Penn State Department of Geography Coffee Hour Colloquium Series*. University Park, PA: March 18.

Roth R.E. 2010. Harrisburg Crime Viz: Geovisual analytics for crime analysis. *City of Harrisburg Bureau of Police*. Harrisburg, PA: April 27.
http:// www.personal.psu.edu/rer198/presentations/Roth_2011_HBP.pdf

**Scholarly Journal Citations**
Based on Google Scholar: 39 thus far (30 by other authors + 9 by paper author/coauthors)

**Other Reports**
Roth RE, F Hardisty, and AM MacEachren. (accepted on extended abstract March 2011) Coordinated, multi-view geovisualization for spatiotemporal crime analysis. In: M Leitner (ed) Crime Modeling and Mapping. Springer, Geotechnologies and Environment Series

**Student Thesis**
Chen, J. 2010: Exploratory Learning From Space-Attribute Aggregated Data –
*A Geovisual Analytics Approach*. Penn State University Department of Geography, Ph.D., Jan. 2010.

Kevin Ross, *SQSynC: Spatial Queries In Synchronous Collaboration*, Penn State, Geography, MS, August, 2010

**Research Assistants**
Robert Roth, Penn State, Geography (PhD candidate)
Alexander Savelyev, Penn State, Geography (PhD candidate)

Xiao Zhang, Penn State, Computer Science (PhD candidate)
Anuj Jaiswal, Penn State, Information Sciences and Technology (PhD candidate)

Graduate Students affiliated with VACCINE but not directly supported during the 2010-2011 project year:

Wei Luo, Penn State, Geography (PhD candidate)
Michael Stryker, Penn State, Geography (PhD candidate)
Bo Yu, Penn State, Information Sciences and Technology (PhD candidate)

Number of Undergraduate Students Supported
Undergraduate Interns (unpaid):
Matthew Murdock, Penn State, Geography
Nicholas Maziekas, Penn State, Geography

**Technology Transitions**
CrimeViz is in the process of being transitioned for use by the Harrisburg Bureau of Police. As part of SP-14, CrimeViz software code has been released (under a LGPL open source license) to colleagues at Purdue as part of the Public Safety Consortium activities and discussion has been initiated with the Justice Institute of British Columbia about the potential for CrimeViz to be used in their ExPod. On the latter, the JIBC was interested in developing a training scenario that uses CrimeViz (subject to available resources to support that activity).
The Health GeoJunction application platform developed in this MDRP was used as the basis for a complementary software application, GrantsForager, funded through a contract from the National Science Foundation as part of Alan MacEachren's participation in the NSF CISE and SBE AC Subcommittee on Discovery in a Research Portfolio: Tools for Structuring, Analyzing, Visualizing and Interacting with Proposal and Award Portfolios.

The SensePlace2 application platform developed in this MDRP was leveraged to acquire option-year funding on an NGA-NURI grant that is focused on finding, extracting, and mapping statements about movement from open source documents (e.g., news stories).

**Follow-Up on Funding/Related Projects**
**Visual Analytics Support for Research Document Retrieval and Analysis**
National Science Foundation (contract) - $32,965
PI: MacEachren
02/01/2010 – 06/30/2011

**Geographic Contextualization for Accounts of Movement (GeoCAM): Representing, Extracting, Mapping, and Interpreting Movement References in Text**
National Geospatial-intelligence Agency - $150,038
PI: MacEachren
08/23/2010 – 8/22/2011 – option year for grant initially funded 8/23/2007
{the focus of this option year is to extend methods for identifying, extracting, and mapping statements about movement in open source documents to apply to documents related to navigation and to flooding; NGA supported our proposal to use the SensePlace2 platform developed in this MDRP as the vehicle for query and display of the documents processed. The

focus on documents about flooding relates directly to the DHS US-Sweden Workshop cited above.}

**RAPID: Text Message-Based Infrastructure for Emergency Response**
National Science Foundation - $75,000
co-PI: Mitra (PI: Yen)
03/15/2010-02/29/2012
Infrastructure for short messages in social media and SMS text for emergency response. Case study with messages obtained from Twitter and Ushahidi related to the Haiti earthquake. Automatic text classification algorithms investigated.

This is a pending proposal:
New Methods for Representing and Interacting with Qualitative Geographic Information
U.S. Army Corps of Engineers - $120,642 pending
PI: MacEachren
{This proposal is to develop and test new text visualization methods that are integrated with interactive maps to represent qualitative geographic information extracted from documents and other media. SensePlace2 has been proposed as the application platform into which the methods and technology developed will be integrated. The goal is to develop methods to support both crisis management and security applications.}

**Other Outcomes/Impacts**
*MDRP 3 & SP-14*
All research within MDRP 3 has been directed to development of visual analytical methods designed to enable crime analysts and investigators to leverage the location and time components of crime and other data that they work with. All resources allocated to MDRP & SP for 2010-2011 were focused on development, assessment, and deployment of CrimeViz. Below, we introduce CrimeViz, summarize research and development progress during the year, and document SP-14 collaborations and technology transition activities. We also highlight complementary GeoViz Toolkit advances that, this year, were supported by sources outside of VACCINE but yielded several publications and tools that cite VACCINE.

*CrimeViz:*
Despite the potential for GIScience techniques to identify and explicate clusters and trends in crime, many police departments lack adequate analytical tools and training to explore and make sense of their crime incident datasets. This concern is particularly poignant in medium to small municipalities that are unlikely to have dedicated crime analysts on staff. Here, we introduce *GeoVISTA CrimeViz*, a web-based map application that supports spatiotemporal exploration and sensemaking of criminal activity. The *GeoVISTA CrimeViz* concept provides understaffed departments with an extensible, easy-to-use tool for conducting spatiotemporal crime analysis and mapping. We are currently in the process of completing a user-centered design approach to ensure that the tool 'works' for its intended audience; a description of this approach and work in progress is provided in the following subsections.

The GeoVISTA CrimeViz application includes a central interactive map (using the Google Maps web mapping service), filtering by crime type, linear and composite animations with VCR controls, an interactive temporal legend that doubles as a frequency histogram, and a set of togglable reference map layers. We are taking a user-centered approach to guide development of the GeoVISTA CrimeViz concept. This approach is following an adaptation of a model for user-centered design developed in previous work by some of our team [Robinson, 2005 #12018].

The revised approach is compared to the more common approach below (figure 1). The key difference is that the revised process starts with prototyping and uses the prototype as a prompt during work domain analysis and related knowledge elicitation activities. As we report in [Roth, 2010 #16928], there are several reasons why it is not always practical to start with a work domain analysis, the most important is that (when users are professionals) the project team may not have access to the target users at the time of development. In addition, when novel methods are being proposed, the goal is not simply to facilitate the current work practice but (potentially) to transform that practice to leverage expertise and data in new ways.
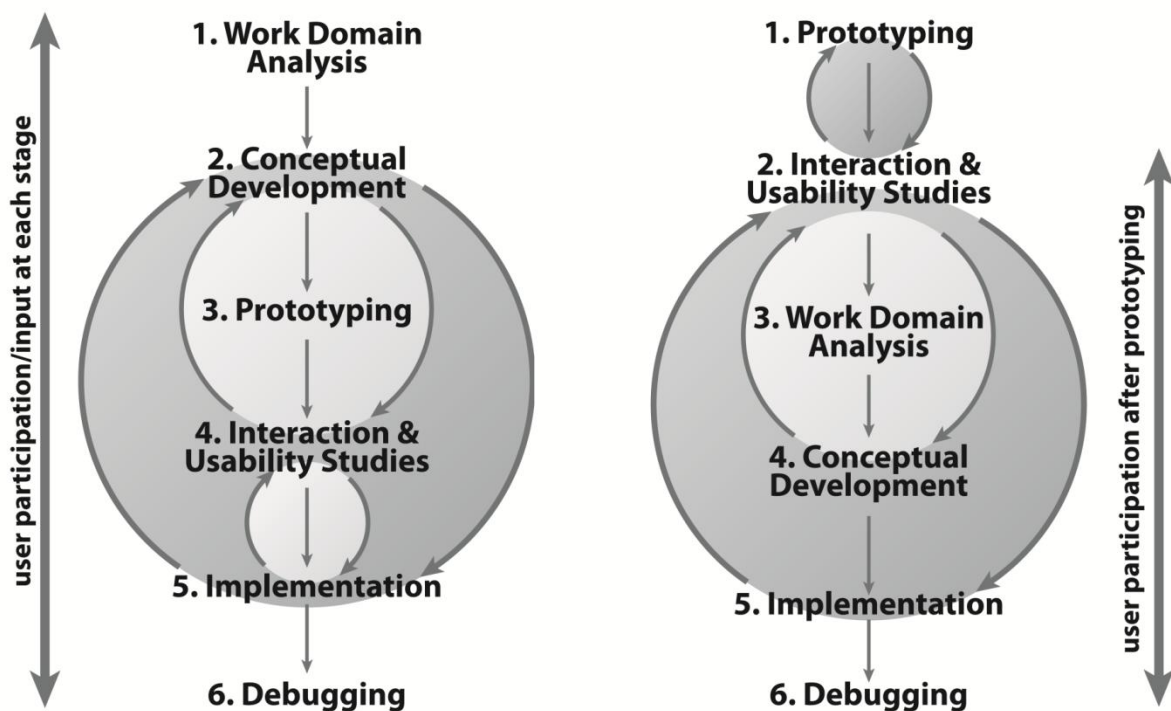


**Figure 2. A model for user centered design developed by Robinson, et al (2005) (left), compared to the user-center design approach developed by Robert Roth as part of this research and adopted for use here (right).**

An initial usability assessment of version one of the CrimeViz applications revealed major interface and mapping problems. The results were used to make substantial improvements to the DC CrimeViz prototype (reported on last year). Our next step in the process was to gather more information about current practices in crime mapping and analysis through interviews, surveys, and hands-on interactive sessions with law enforcement personnel. This included in-depth, one-on-one interviews (with 9 analysts / decision-makers at 6 municipal and 1 federal law

enforcement agency) to assess the current crime analysis practice of law enforcement agencies, focusing on currently met and unmet needs. We obtained information on: (a) characteristics of data used and desired, (b) kinds of visual representations used and desired, (c) kind of tool interactivity in practice now and the limitations that both representations and interaction methods pose for analysis, (d) current use of spatial analysis methods if any, (f) current use of temporal analysis methods if any, and (g) ways in which maps are used and impediments to use. As a follow up, we asked professional crime analysts and decision-makers to test out the tool on their own and respond to a series of Likert-based questions (reactions were positive with 70% strongly agreeing that CrimeViz was easy to use and no participants indicating that they would need support of a technical person to use the tools). However, some concerns were raised about missing functions and the match to the type of work each participant does now (the modal response remained positive on these factors).
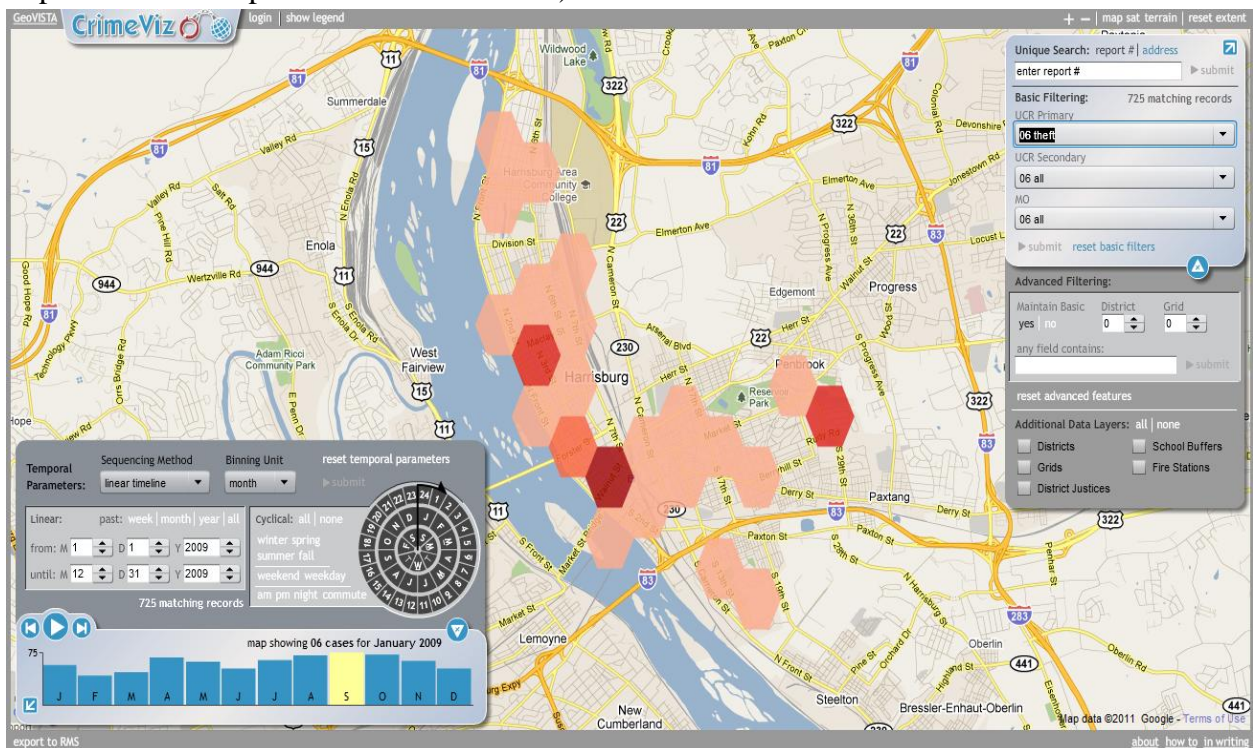


**Figure 3. GeoVISTA CrimeViz allows crime analysts to generate both linear (e.g.,. 2008, 2009, 2010) and composite (e.g., average year by month, average week by day) views of time. The figure shows a clear seasonal pattern when examining incidents of theft by month, with peaks occurring in the early Spring and late Summer/early Fall and lows occurring during the middle of Winter.**

Feedback from these activities has been integrated into version two illustrated above (figure 3), which is being designed specifically to support needs of the Harrisburg Bureau of Police. The revised GeoVISTA CrimeViz application features an interface that simplifies complex queries using combinations of URC codes, MOs, and keywords and also implements a robust approach to dealing with the temporal component of the data, allowing for filtering by time range, hour of the day, day of the week, and month(s) of the year. Time ranges can be explored independently

or time series can be explored through animation. GeoVISTA CrimeViz also supports analysis of "composite time"; for example, by grouping multiple years' data into an average week-by-day or year-by-month, the user can uncover patterns. The system deals with data scaling issues through automated spatial aggregation to a regular polygon tessellation (or any user-supplied set of polygon boundaries).

As noted above, a graduate student in the Master in Criminal Justice program at Penn State Harrisburg (a former police officer) has collaborated with our team through her in-site internship with the Harrisburg Bureau of Police. Additional on-site work by the research team in Harrisburg is planned for April, 2011 to continue transition activities. A comprehensive use and usability assessment of the tools is planned for summer 2011.

SP-14 focused on coordination with Purdue and the Justice Institute of British Columbia about the potential to use CrimeViz in training exercises provided through the JIBC ExPod. The software code for CrimeViz (with LGPL open source license) has been shared with collaborators at Purdue. Remote meetings have been held with personnel at JIBC to demonstrate CrimeViz and to discuss how it might be used in their training activities. The JIBC did not think it was feasible to integrate CrimeViz with their X-Pod tool (due, in part, to lack of resources and in part to differences in technology). However, a need was identified to include a crime mapping component in one of their training scenarios. They indicated that they may want to use a synthetic dataset for a hypothetical town in British Columbia. They also indicated that they may be interested is using a Harrisburg-based scenario with public data, if PSU developed one. The latter would, of course, require new VACCINE resources to accomplish. But, we also have interest from the Director of the Graduate Certificate in Geospatial Intelligence to use an exercise build around CrimeViz within a course in that program should such an exercise be created.

*GeoViz Toolkit:*
In addition to CrimeViz, advances in spatio-temporal analysis methods have been achieved and integrated into the GeoViz Toolkit. The GeoViz Toolkit provides a desktop complement to CrimeViz and includes integrated visual-spatial statistical methods. Due to budget reductions, however, none of the GeoViz Toolkit advances this year were funded through VACCINE; progress continued with funding from the Gates Foundation and from NGA (with the application domain in both cases being infectious disease and terrorist events, respectively).

*MDRP 13 & SP-15*
In this section, brief updates on Health GeoJunction and SensePlace (both reported on last year with publications and follow on work funded through this year) are provided. Then, research on the Context Discovery Application is summarized briefly. This represents the conclusion of a large project carried out across our NEVAC (an RVAC) and VACCINE funding and with additional support from NSF; with one key paper published this year and another invited for a special issue. Then, our current work on SensePlace2 (a sub-project and software application introduced in last year's report) is summarized.

*Health GeoJunction:*

Research on Health GeoJunction was completed in 2009-2010 and reported on in last year's report. This year saw publication of a paper documenting that research (that yielded a "Highly Accessed" rating in the International Journal of Health Geographics; http://www.ij-healthgeographics.com/content/9/1/23 as well as 17th on their list of "most forwarded" for the year: http://www.ij-healthgeographics.com/mostforwarded/year), multiple demonstrations of tool capabilities, and follow up research on SensePlace2 which leveraged selected ideas and some of the software code to support a broader range of place-time-concept filtering of document-derived data. In addition, Health GeoJunction was the source of ideas that leveraged NSF support to develop GrantsForager; this project demonstrated the generality of the approach. Health GeoJunction, itself, has not been developed further during the year, thus no additional details are provided here. However, the GrantsForager work for NSF resulted in a report on that effort that is included as an appendix to this report.

*SensePlace:*

Research on SensePlace was reported on initially in the 2009-2010 report (within the MDRP 1 report). It is cited here because core ideas developed and implemented were adapted within SensePlace2 and the paper documenting development and application of SensePlace that was submitted last year was revised, accepted, and published this year. SensePlace is an outgrowth of an earlier effort, the Context Discovery Application (CDA), begun within NEVAC (our Regional Visualization and Analytics Center that preceded VACCINE). Brian Tomaszewski (now on the RIT faculty) continued work on CDA in collaboration with Alan MacEachren. Results of that collaboration appeared in the IEEE VAST Proceedings published in this project year and the paper was invited for submission (in extended form) for a special issue of Information Visualization featuring several of the top VAST papers. We summarize that work here.

*Context Discovery Application:*

CDA is a geovisual analytics environment focused on document foraging and sensemaking. The application is the outgrowth of research to develop a theoretical model of geo-historical context illustrated in the figure at right. The GHC model (see figure 3) serves two purposes. *First*, it provides a formal structure for the theoretical and conceptual components of GHC; these include (a) events, places, and concepts as well as (b) relationships and constraints among these components such as scale, and spatial and temporal topological relationships. The GHC model is specifically structured using three sub-models - geographical, historical, and conceptual, which represent windows into locality of context, along with compatibility relations among the components. Special characteristics of the GHC model when compared with other spatiotemporal models include the emphasis on (a) modeling context in particular, (b) modeling derived knowledge and information rather than raw observational data, and (c) modeling reasoning contexts using local model semantics. *Second*, the model can be used as a conceptual template for structuring and representing specific information instances retrieved, compiled, developed, and ultimately used as part of foraging and sensemaking processes. Context

information can then be applied to fulfill a task or achieve a goal requiring geographical, historical, and thematic interpretations of situational factors. The GHC model is represented formally through an OWL-DL computational ontological structure (one of several sublanguages of the Web Ontology Language) which is effective at representing, capturing and describing aspects of real-world contextual information in computer readable formats. As reported in the VAST 2010 paper, CDA usability and utility was assessed. United Nations staff participated in the utility evaluation and provided input on the potential of the GHC framework and its



**Figure 4. The overall process by which GHC is produced.**

instantiation in the CDA to support work in humanitarian crises management. Tomaszewski (funded separately) is currently working with the UN to leverage methods and software components incorporated within CDA for UN web mapping applications.
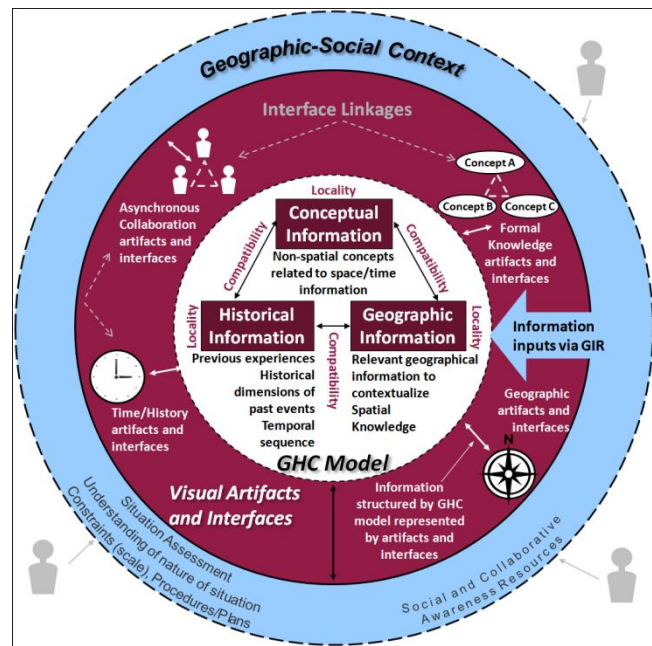
*SensePlace2:*

Social media are becoming increasingly geographic. Following from this trend, maps that depict a wide range of data from geo-located social media are becoming fairly common. Most social media mapping research and most web maps of social media data thus far focus on the tasks of capturing, locating, and displaying data extracted from social media. This task itself is a challenge; for Twitter alone, the number of tweets reached 110 million per day in January, 2011 and that number is escalating rapidly (source: Forbes.com blog: http://tiny.cc/s8pbc). However, extracting meaning from these large, messy, but potentially important data sources is an even greater challenge that is just beginning to be addressed. One application domain for which the need is particularly important is crisis management; the potential usefulness of social media in this domain is reflected in the increased attention given to social media by major crisis management organizations (e.g., the Red Cross Emergency Social Data Summit held in August, 2010) and by those in the crisis management research community.

In this research, the focus is on efforts to address the challenge of social media information mapping and analysis more systematically, with an emphasis on potential applications for crisis management. More specifically, we apply a visual analytics perspective to develop and implement visually-enabled information foraging and sensemaking tools for leveraging data made publically available in social media. Initial efforts are focused on the rapidly growing resource of Twitter tweets, both those that include geographic location and those for which location can be inferred through specific hashtags or by automated entity extraction methods. A primary contention that underlies our work is that a visual analytics approach is particularly relevant for dealing with Twitter because of the scale of data involved, the limited details in each

tweet combined with potential to generate insights by combining many small fragments of information, plus the challenges in identifying both places and topics in short, cryptic tweets. No computational methods will be completely successful in tweet retrieval and interpretation. Thus, a tight integration between computational and visual methods is needed; the computational methods are essential to handle the scale of data and the visual methods are essential to help users deal with uncertainty about information quality and relevance.

The research is following a user-centered design / cognitive systems engineering approach in which knowledge and insights from practitioners who are potential users of the methods and tools are obtained systematically and repeatedly throughout the development process. During 2010-2011 periods, we completed a first step in this process. Specifically we surveyed crisis/emergency management professionals who responded to questions about potential uses of social media in crisis management and provided feedback on an early design mockup for SensePlace2. The survey had two primary goals: two primary goals; (1) to develop a clear understanding of how emergency management professionals currently use social media tools and (2) to elicit what emergency management professionals envision doing with social media geovisual analytics tools in the future.

To address these goals, we developed a 21 question survey delivered on-line using SurveyMonkey. To recruit participants, we posted a request to the International Association of Emergency Managers (IAEM) listserv and to firstresponder.gov. Our request yielded 46 responses. Because we did not require participants to answer all questions, some received less than 46 responses. The first question in our survey asked respondents to characterize their work experience in emergency management. Respondents had a high level of expertise overall; the majority (53.7%) had over 9 years of experience, while 22% had between 6-9 years, 2.4% between 3-6 years, 12.2% between 1-3 years, and 9.8% claimed no experience working in emergency management.

Initial results of the survey are summarized in a web document that accompanies the paper submitted to VAST about this research thus will not be detailed here; the summary is available at: http://www.geovista.psu.edu/SensePlace2/VAST_2011_Appendix_A.pdf. Overall, the results of our survey suggest that those in the emergency management domain are actively using social media tools, that they expect to be able to explore multiple kinds of geography associated with social media, and that they anticipate making use of this information primarily for situational awareness activities. Our results provide a roadmap for future design iterations of SensePlace2 that include more useful mapping tools, data interoperability features, and analytical reporting capabilities.

Input obtained from the survey cited above is being integrated into design and refinement of SensePlace2 and we plan follow up knowledge elicitation activities with practitioners along with empirical evaluations of SensePlace2 over the coming year. The default SensePlace2 interface at the end of the 2010-2011 project year is described below. It includes a query window, map,

time-plot / control, relevance-ranked list of tweets, and task list. The primary display views (map, time-plot/control, and tweet list) are dynamically coordinated. Each view is introduced below; then cross-view linking is discussed.

*Query window*: Users can enter single or multi-term queries and these can include place names. Each query retrieves a new set of information that is processed to populate display views.

*Map*: The map provides both overview, in the form of a gridded density surface representing all tweets that match the query, and detail in the form of point-based depiction of the most relevant 500 tweets. The density surface is generated for the globe and currently depicts frequency counts for tweets aggregated to 2 degree grid cells (grid resolution is flexible, but that flexibility has not yet been made accessible to users). A quantile classification scheme is applied, to allow comparison from one query to the next, and a sequential color scheme is used with dark=highest. It is likely that some locations for the top 500 tweets can have multiple tweets, thus those locations are depicted with range-graded sizes for 1, 2-5, and >5 tweets from/about a place.

*Time-filter and frequency display*: The time-filter and frequency display doubles as a compact representation of the frequency distribution of tweets that match the query across the full time span of data in the database and a control to filter tweets by time. Specifically, both ends of the time range selector can be dragged and once a time range is set, that range can be shifted along the timeline by clicking on the time snap bar and dragging.

*Tweet list*: The 500 most relevant tweets are displayed in a scrolling list. The list can be sorted by relevance (the default), time, and place. Hierarchical sorting is also enabled (e.g., with time as the primary sort and relevance as the secondary to highlight recent and relevant tweets).

*Task list*: The task list view (not yet fully implemented) will allow users to label results of a query and store that result in a history. These stored queries will retain user set parameters that include any place and time filtering as well as decisions to promote at tweet to high relevance or hide an irrelevant tweet.

Our initial focus has been on leveraging the *explicit geographic information* in tweets (when users add place or coordinate information to tweets), the *explicit (but imprecise) geographic information* contained in the user's profile (their time zone), and the *implicit geographic information* that can be inferred through identification of place-related hashtags or automated entity extraction methods. An example view from the interface is below (Figure 4).
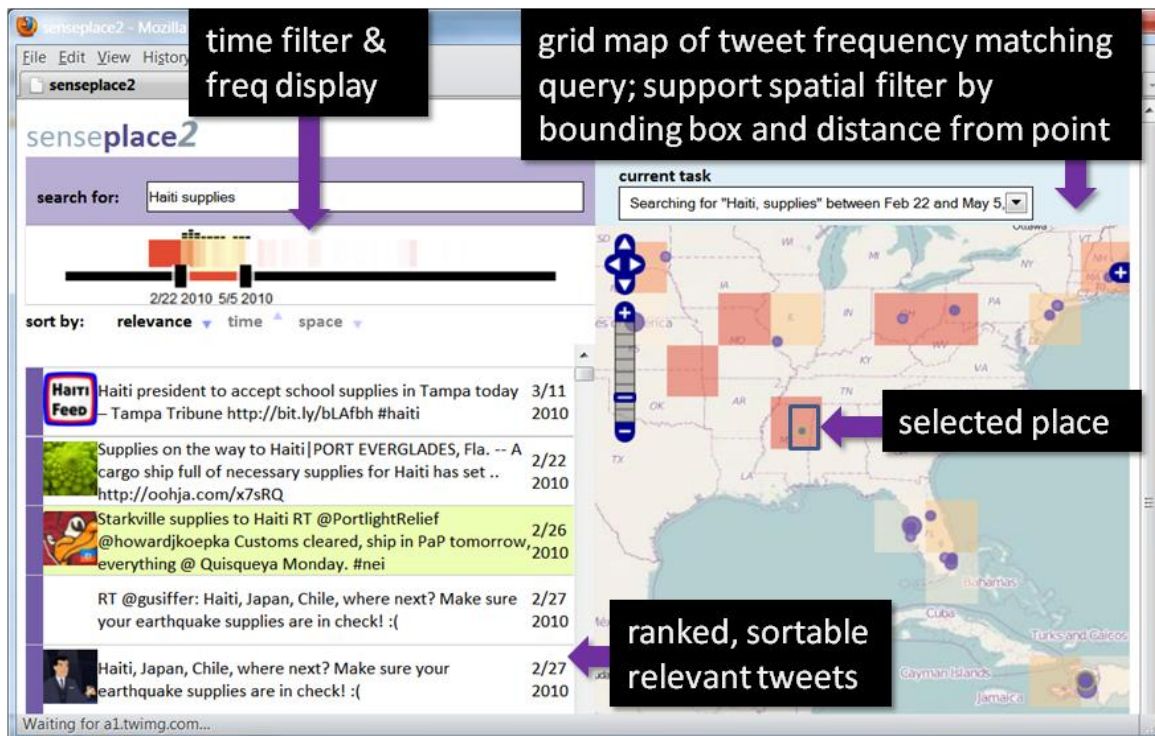
**Figure 5. SensePlace2 interface: time constraint on query for 'Haiti supplies' plus spatial selection of relevant results.**

The overall system architecture consists of a central database that contains data tables, associated named entity tables, location tables, and text search indices for each information channel that support both the text processing tools and the UI. The architecture is designed such that a wide variety of information channels can be easily integrated and allow the user interface/ processing software to quickly access and process this information; thus the design plans for moving beyond twitter as the only information input. Figure 5 illustrates the current Twitter Information channel storage and processing backend. Each channel of information that is currently stored consists of four backend processing daemons:

*Channel Search/ Data Load Daemon*: Each channel has an associated daemon that collects information of interest. For example, the twitter channel search daemon currently uses a table containing queries that are passed on to twitter using the Twitter Search API (http://search.twitter.com/api/). Twitter then returns a set of tweets with auxiliary metadata (tweet geolocation, time of creation, user id etc.) which are then stored in the channel data table. Each channel data table typically contains a single column of text (in this case it contains the text contained within the tweets returned from a twitter search) which is of interest. This column information is stored in the system information table as it will be used by the text processing software.
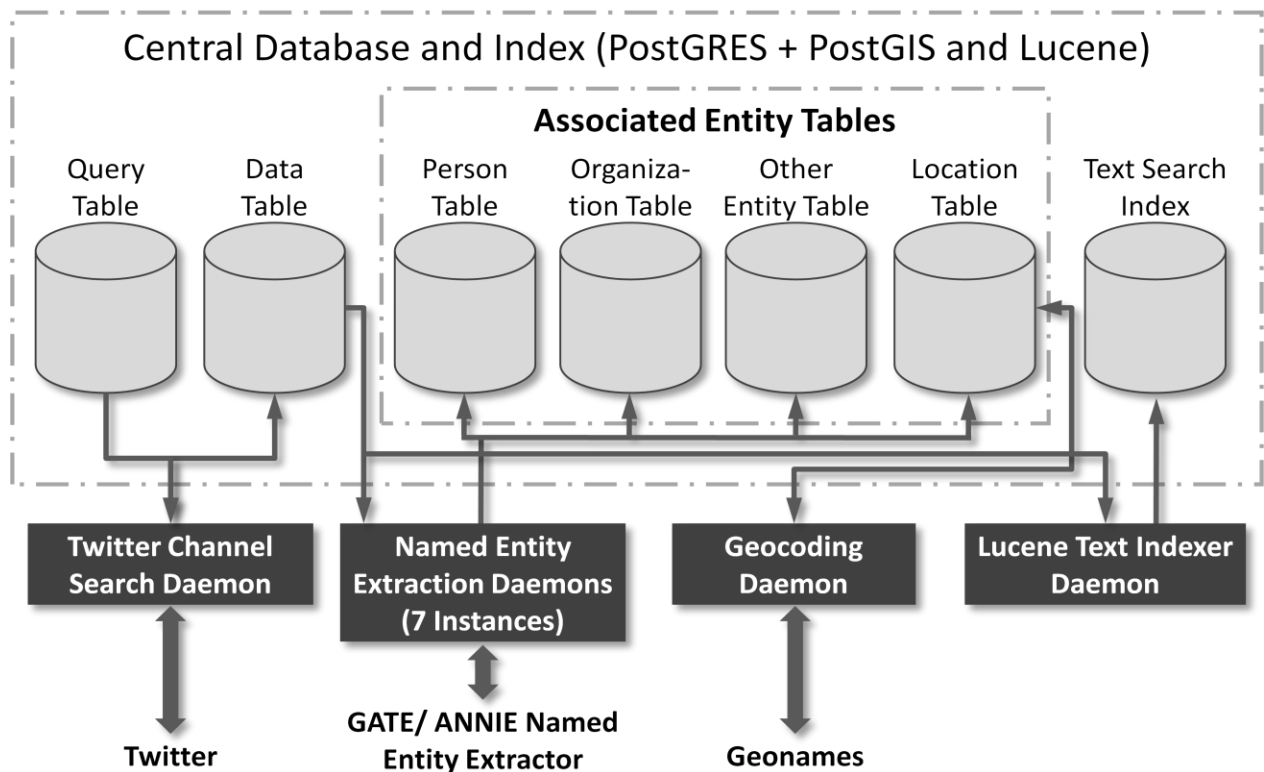
**Figure 6. Twitter information channel storage and processing backend.**

*Named Entity Extraction Daemon*: The named entity extraction daemon extracts the named-entities contained within the text present for each channel and writes this information to the associated named-entity tables in the database. There are four broad categories of named entities extracted, namely people, locations, organizations and others. The current daemon uses a custom developed ANNIE named-entity extractor that is based on GATE (http://gate.ac.uk/ie/). The named-entity extraction daemon is a distributed application that runs on multiple machines on the same network allowing a relatively large volume of text to be processed in real time. The twitter channel currently uses 7 machines to perform named entity extraction.

*Geocoding Daemon*: Once the named-entity extraction has occurred, a location table is generated that contains text information (e.g. "Haiti") regarding locations contained within a channel. While this is suitable for text querying, such information is insufficient to perform spatial querying on this channel, e.g., to find the closest tweets to a location. The Geocoding daemon processes the location table (for all channels) and geocodes these locations with the point coordinates of the best matching location returned from Geonames.org web service. These point coordinates are then stored and indexed within the location table allowing for quick generation of heat maps and retrieval of relevant tweets in a desired location neighborhood.

*Lucene Text Indexer Daemon*: Each channel typically contains a single text column in the channel data table. For example, for the twitter channel the data table contains one column that stores all tweets that were extracted by the channel search daemon. The Lucene text indexer daemon creates a fielded Lucene text index for the channel text for full text querying as well as for more advanced retrieval of relevant tweets within a date range. The main fields are text,

50

location, organization, people and other entities; boosting of selected fields (e.g., location) is used to improve relevance ranks.

The initial version of SensePlace2 demonstrates the potential of using a place-time-concept framework to support information foraging and sensemaking in this domain. The scenario-based design / cognitive systems engineering approach applied to development provides an efficient strategy to refine design and a framework for structured user studies that will guide that design.

**Collaboration with other VACCINE project teams, National Labs, and other Homeland Security Centers of Excellence (including CREATE, FAZD, NCFPD, START, PACER, CAMRA, ALERT, NCBSI, MIREES, NDCIEM and NTSCOE):**
Cross-University collaboration is underway with Purdue for CrimeViz, for which (as noted above) software code has been provided to Purdue to support collaborative development and distribution through the Public Safety Consortium and (potentially through the JIBC ExPod).
We have been invited to attend the DHS Workshop on *A Visualization and Analytics Approach to Flooding and Pandemics* (to be held 27-28 April 2011).

We have discussed plans to leverage our research on geographically-enabled social media analytics with other members of the Theme 3 VASA (Visual Analytics for Security Applications) team (pending anticipated funding to participate actively in that effort).

We have initiated discussion with NVAC at PNNL about coordination on (geo)visual analytics applied to social media content to support crisis management and other homeland security applications.

**School**
Purdue University

a) **Project Name/Theme**
MDRP 18: Visual Analytics for the DHS Centers of Excellence

Theme 2: Federal Operating Component Projects

b) **PI/Co-PI and Institution**
Niklas Elmqvist, Purdue University

**Collaborating Partners**
David Ebert (VACCINE/Purdue), John Stasko (VACCINE/Georgia Tech), Brian Fisher (VACCINE/SFU), Ed Hovy (CCICADA/USC)

**Collaborating End-Users**
DHS University Programs

c) **Research Problem (Abstract)**
The DHS Centers of Excellence has produced a wealth of research, education materials, and projects over the years, and this material will continue to grow as the activities of the various centers expand. The ability for reflection, analysis, and investigation of this material is vital, and may yield both new and old connections, ideas, and thoughts. Furthermore, exposing the material to the scientific community and other government agencies, as well as the general public, is an important goal of our center. We leverage visual analytics methods for interacting and exploring this large knowledge resource in an interactive visual form that is easy to access for quick browsing, yet powerful enough to support advanced analysis and review.

This project develops table-based, graph-based, and timeline-based visualizations of the DHS Centers of Excellence program in an easily accessible format, but which also supports faceted browsing and querying on the multiple dimensions associated with the data. These visualizations will be easy to expand with future data without requiring redesign or rebuilding.

d) **Technical Approach**
Data Collection—Compiling & Sorting DB; Analytic—Case Studies; Nature of Research—Hybrid Basic-Applied

e) **Deliverables**
Initial prototype has been implemented and is being deployed on the VACCINE web site. Data collection is well underway.
https://engineering.purdue.edu/~elm/projects/coe-explorer/

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
S. Ghani, N. Elmqvist. Improving Revisitation in Graphs through Static Spatial Features. In *Proceedings of Graphics Interface*, to appear, 2011.
https://engineering.purdue.edu/~elm/projects/ssgf/ssgf.pdf

S. Ghani, N. Henry Riche, <u>N. Elmqvist.</u> Dynamic Insets for Context-Aware Graph Navigation. *ComputerGraphics Forum* (IEEE EuroVis 2011 proceedings), 30(3):861–870 2011. [54/190, 28% acc. rate]
https://engineering.purdue.edu/~elm/projects/dyninsets/dyninsets.pdf

K. Kim, S. Ko, <u>N. Elmqvist</u>, D. Ebert.  WordBridge: Using Composite Tag Clouds in Node-Link Diagrams for Visualizing Content and Relations in Text Corpora.  In *Proceedings of the Hawaii International Conference on System Sciences* (Visual Analytics Minitrack), 2011.
http://engineering.purdue.edu/~elm/projects/wordbridge/wordbridge.pdf

**Presentations**
KyungTae Kim.  WordBridge: Using Composite Tag Clouds in Node-Link Diagrams for Visualizing Content and Relations in Text Corpora.  Presented at the *Hawaii International Conference on System Sciences* (Visual Analytics Minitrack), Kauai, Hawaii, January 2011.

Niklas Elmqvist.  Dynamic Insets for Context-Aware Graph Navigation.  Presented at *IEEE EuroVis 2011,* Bergen, Norway, June 2011.
http://www.slideshare.net/NickElm/dynamic-insets-for-contextaware-graph-navigation

**Patents/Copyrights**
M. Boutin, S. Biddanda, N. Elmqvist, D. Ebert. Method for Visualizing the Results of an Electronic Search. Provisional patent 61/348,542, May 2010.

**Research Assistants**
Sohaib Ghani, Ph.D. Student (3$^{rd}$ year), School of Electrical and Computer Engineering, Purdue University.
Number of Undergraduate Students Supported--(2 SURF students in Summer 2010).

**Other Outcomes/Impacts**
The current work on the COE Explorer focuses around the following four areas:

**Interactive social network analysis** (Elmqvist – Purdue) We are working on improving the navigation and interaction aspects of the COE Explorer application to make it possible to perform some social network analysis within the tool.  The screenshots below show the COE Explorer tool using an attribute-based layout on the degree centrality of authors in the OUP network.

**Improved dataset** (Hovy – USC) – Based on the existing OUP database, our emphasis is on improving the quality and the contents of the dataset being visualized using the COE Explorer. So far, there has been progress on extracting descriptive keywords from the scientific literature for the CCICADA and VACCINE centers, and additional COE publications are currently being analyzed to be included into the dataset.

**Managing scale** (Elmqvist, Ebert – Purdue) – For the tool to support the new and larger datasets, we are working on integrating hierarchical aggregation so that the network being visualized can be automatically summarized.  This would allow it to fit on comfortably on the screen both for a

desktop computer and for a mobile device, and would also reduce the cognitive demands of navigating in a massive network.

Interaction **design and presentation** (Fisher - SFU) – We are working on improving the user interface and information presentation of the COE Explorer.  One aspect of this is to also generate an instructional video that showcases the utility of the Explorer for different analytical and presentation tasks.

**School**
Purdue University

a) **Project Name/Theme**
Investigation of the Impacts of Visualization Literacy on the Performance of Visual Analysis

Theme 6: International Collaborations (also applicable to work on VASA)

b) **PI/Co-PI and Institution**
Ji Soo Yi, School of Industrial Engineering, Purdue University

**Collaborating Partners**
Brian Fisher, School of Interactive Arts and Technology, Simon Fraser University

**Collaborating End-Users**
Local Law enforcement

c) **Research Problem (Abstract)**
The goal of this project is better understand the roles and influences of "visualization literacy" in the context of investigative analysis. We hypothesized that background knowledge in how to interpret and use various visualization techniques influence the performance of investigative analysts while using visual analytic tools, such as Jigsaw and INSPIRE. However, since we do not have a clear definition of visualization literacy at this moment, we designed and conducted an exploratory and qualitative study to understand visualization literacy. More specifically, we observe how research participants conduct investigative analyses with a visual analytic tool and attempted to identify "roadblocks" that prevented them from making full use of such a visual tool. The results show that there are different types of roadblocks that people run into. One interesting finding is that participants have troubles in fully utilizing visualization techniques even though they get through comprehensive introduction and report that they understand those visualization techniques clearly. They seem to be a gap between knowing and using a visualization technique. We also observed professional investigative analysts using the experimental setting, so expect to better understand the relevant issues.

d) **Technical Approach**
Data Collection—Survey, Pair Analytic Method; Analytic—Grounded Theory; Nature of Research—Basic

e) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Kwon, B. C., Fisher, B., & Yi, J. S. (n.d.). Visual Analytic Roadblocks for Novice Investigators, *VAST 2011*.
(conditionally accepted)

**Research Assistants**

Bum chul Kwon, School of Industrial Engineering, Purdue University

**Other Outcomes/Impacts**

Identifying the five visual analytic roadblocks and better understanding the nature of these issues through qualitative and in-depth pair analytic methods is the major accomplishments of this seed project. The results of this research were reported to one of major conferences in the field, The IEEE Conference on Visual Analytics Science and Technology, and the paper was well accepted by the reviewers.

**School**
Purdue University

a) **Project Name/Theme**
MDRP 6: Video Surveillance Visual Analytics

Theme 1: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**
Edward J. Delp, Purdue University

**Collaborating End-Users**
Purdue University Police Department

c) **Research Problem (Abstract)**
The goal of this project is to detect activities, anomalies, and other events of interest using visual methods. Two tasks have been described in this report. The first task is to infer behavioral information from the trajectory of an object. The object's position is obtained by tracking and is mapped to true ground positions using camera calibration. We then analyze the patterns in the velocity and the shape of the trajectory to determine anomalous motion. This analysis can help detect potential threats in vehicles and persons for safety and infrastructure protection. The second task is to automatically and accurately estimate the pedestrian "flow" from a monocular video. The term "flow" is defined as the number of persons passing a designated region (referred to as trip wire) in a unit time. This analysis can help determine the average pedestrian traffic at a point of interest, detect important ingress/egress points and possible bottlenecks, and also detect panic situations when the crowd motion patterns become anomalous.

d) **Technical Approach**
Data Collection—Data Mining, Expert Consultation, Field Monitors; Analytic—Modeling, Statistical Analysis; Nature of Research—Applied Basic

e) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Satyam Srivastava, Ka Ki Ng, and Edward J. Delp, "Crowd Flow Estimation Using Multiple Visual Features for Scenes with Changing Crowd Densities," *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance*, Klagenfurt, Austria, August-September 2011

Satyam Srivastava, Ka Ki Ng, and Edward J. Delp, "Co-Ordinate Mapping and Analysis of Vehicle Trajectory for Anomaly Detection," *Proceedings of the IEEE International Conference on Multimedia and Expo*, Barcelona, Spain, July 201

Satyam Srivastava, Ka Ki Ng, and Edward J. Delp, "Color Correction for Object Tracking Across Multiple Cameras," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011

Ka Ki Ng and Edward J. Delp, "Background Subtraction Using A Pixel-Wise Adaptive Learning Rate For Object Tracking Initialization," *Proceedings of the IS&T/SPIE Conference on Visual Information Processing and Communication*, San Francisco, California, January 2011

Satyam Srivastava and Edward J. Delp, "Standoff Video Analysis for the Detection of Security Anomalies in Vehicles," *Proceedings of the IEEE Applied Imagery Pattern Recognition Workshop*, Washington, D.C., October 2010

**Research Assistants**
Ka Ki Ng, Purdue University, School of Computer and Electrical Engineering

**Other Outcomes/Impacts**
We test our proposed methods using publicly available datasets as well as on a dataset provided by the Purdue Police Department and good accuracies are obtained. Experimental results show that our proposed method on flow estimation works well even when the crowd density levels change significantly. Similarly our method on motion analysis was found to work well even under difficult and uncontrolled outdoor conditions.

**School**

Purdue University

a) **Project Name/Theme**

Visual Analytics Decision Support Environment for Epidemic Modeling and Response Evaluation

Theme 1: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**

David Ebert, Ross Maciejewski, Purdue University

**Collaborating Partners**

Foreign Animal and Zoonotic Disease Center

c) **Research Problem (Abstract)**

In modeling infectious diseases, scientists are studying the mechanisms by which diseases spread, predicting the future course of the outbreak, and evaluating strategies applied to control an epidemic. While recent work has focused on accurately modeling disease spread, little work has been performed in developing interactive decision support tools for analyzing the epidemic catastrophes and evaluating potential disease mitigation strategies. The absence of such tools makes it difficult for researchers, analysts and public health officials to evaluate response measures within outbreak scenarios. As such, our research focuses on the development of an interactive decision support environment in which users can explore epidemic models, mitigation strategies and the impact of both. This environment provides a spatiotemporal view where users can interactively utilize mitigative response measures, such as quarantine and transportation controls and observe the impact of their decision over time. Our system also provides users with doubly linked decision history visualization and navigation tools that support the simultaneous comparison of mortality and infection rates corresponding to different response measures at different points in time. The architecture of our system is flexible enough to model variety of epidemic catastrophes by integrating corresponding epidemiological models. We use hypothetical scenarios and dataset to perform our case studies. As the user inserts decisions points, scrolls through time, and revisits other scenarios, these interactions are tracked and displayed in the decision history view. Our case studies are the following: Pandemic Influenza with airline transportation and Rift Valley Fever, a mosquito-borne zoonotic disease. The results demonstrate the success of our system in designing effective mitigative strategies.

d) **Technical Approach**
**Data Collection Methods**

Data Collection--Compiling & Sorting DB, Data Mining; Analytic—Modeling, Sampling, Statistical analysis; Nature of Research—Hybrid Basic-Applied

e) **Deliverables**

There are two software products to date from this project. The first system enables visual analytic decision making and evaluation of interdiction measures for Rift Valley fever (RVF), a zoonotic disease that could have a devastating impact on the US cattle industry. The second

system allows visual analytic decision making to respond to and curtail the spread of pandemic influenza.

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Afzal, S., Maciejewski, R., Ebert, D., "Visual Analytics Decision Support Environment for Epidemic Modeling and Response Evaluation," conditionally accepted pending minor revision to *Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST),* 2011.

**Presentations**
March 2011 – Shehzad Afzal, Visual Analytics Decision Support Environment for Epidemic Modeling and Response Evaluation, DHS University Programs Summit

**Research Assistants**
Current Graduate Students *(include name(s), university and department)*
Shehzad Afzal, Purdue University, Computer Science

**Other Outcomes/Impacts**
This project has shown the potential of integrated simulation, visual analysis, and linked interactive spatiotemporal displays with decision space displays for evaluating and optimizing multiple interdiction strategies.

**School**
Purdue University

a) **Project Name/Theme**
Coast Guard Search and Rescue Visual Analytics

Theme 2: Federal Operating Component Projects

b) **PI/Co-PI and Institution**
David Ebert, Ross Maciejewski, Purdue University

**Collaborating Partners**
United States Coastguard

**Collaborating End-Users**
United Stated Coast Guard District 9

c) **Research Problem (Abstract)**
We have developed a comprehensive visual analytics system to analyze the risks to maritime traffic and public safety associated with the reallocation of available Coast Guard resources and shutting down of stations mandated due to budget changes, retiring volunteers, among other factors. The system includes linked views and interactive displays that allow the interactive analysis of trends, patterns and anomalies among the U.S. Coast Guard Search and Rescue (SAR) operations and their associated sorties. Our system allows users to determine the potential increase or decrease in risks associated with shutting down certain station(s) in terms of factors including response times, potential lives and property lost and reallocation of available resources. The system also enables a thorough assessment of all SAR operations conducted by each U.S. Coast Guard station in the Great Lakes region. In addition, the system provides officials with the tools to determine which Coast Guard stations were more optimally suited to assume control of the operations of the closed station(s) by comparing the optimal transportation routes of the available stations to all the SAR cases previously handled by the closed station. This project demonstrates the effectiveness of visual analytics in analyzing risk within the maritime domain. Our system provides analysts with a suite of tools for analyzing risks and consequences of taking major decisions that translate into important measures including potential lives lost and property damaged. Our system is currently being used by the U.S. Coast Guard Ninth District that is responsible for all Coast Guard operations throughout the five U.S. Great Lakes. Our short term future plans include deploying our system to assist in the analysis and optimization of all operations conducted by the Ninth U.S. Coast Guard District and expanding the use of this tool to other USCG Districts.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Data Mining; Analytic—Modeling, Sampling, Statistical analysis; Nature of Research—Hybrid Basic-Applied

e) **Deliverables**
cgSARVA software has been delivered to USCG D9 and Atlantic Area Command Operations Analysis group (LANT-7).

f) **Project Outcomes (Results, Accomplishments, Conclusions)**

**Publications**
Malik, A., Maciejewski, R., Maule, B., Ebert, D., "A Visual Analytics Process for Maritime Resource Allocation and Risk Assessment," conditionally accepted pending minor revisions, *Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2011.

**Presentations**
March 2011 – Abish Malik, Coast Guard Search and Rescue Visual Analytics, DHS University Programs Summit

**Research Assistants**
Current Graduate Students *(include name(s), university and department)*
Abish Malik, Purdue University, Computer Science
Number of Undergraduate Students Supported
Daudi Muhamed, (SURF 2010)

**Follow-on Funding/Related Projects**
This initial work led to additional funding from USCG through the DHS BOA mechanism for visual analytics for Port Security and collaboration with USC CREATE for security patrol scheduling for USCG D1 ($160,000)

**Other Outcomes/Impacts**
This tool has provided guidance and analytical evidence to USCG D9 to evaluate the effectiveness of their resource allocation for their Search and Rescue Mission in the Great Lakes and the role the Coast Guard Auxiliary has in this mission. The tool enables the USCG to consider what-if scenarios for resource allocation, evaluate the effect of the reduction of Auxiliary Stations and reallocation of resources, as well as staffing hours to still successfully fulfill their SAR mission.

**School**
Purdue University

a) **Project Name/Theme**
Visual Analytics Law Enforcement Technology

Theme 1: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**
David Ebert, Ross Maciejewski, Tim Collins, Purdue University

**Collaborating End-Users**
Tippecanoe County Sheriff's Office, Lafayette Police Department, West Lafayette Police Department, Purdue Police Department

c) **Research Problem (Abstract)**
The exploration of criminal incident reports for detecting trends, discovering anomalies and evaluating resource usage is an ever-expanding issue for law enforcement agencies. As such, tools need to be developed that assist law enforcement officials in their analysis in order to take preventive measures and judiciously allocate available resources. In order to better facilitate crime incident analysis, we have developed a comprehensive visual analytics system for both PCs and mobile devices which provides police officials with access to their data on the fly.

The application was developed for both the Windows PC environment and the Apple's iOS platform and can run on any compatible devices including the iPhone, iPod, and iPad. Our system allows users to visualize data geo-spatially on a map and provides filtering tools that filter crime by the type of offense committed. The crime being visualized can additionally be filtered by time to analyze the correlation of different crimes with time. These set of tools enable the users to analyze data more efficiently and accurately wherever they are with limitations only to places without mobile signal coverage. Our system enables users to view a history of previous crime incidents and forecast a pattern of crime using automated algorithms. Moreover, the users have the ability to receive immediate feed of events. This is especially beneficial to first responders and allows them to better deal with any catastrophic situation.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Data Mining; Analytic—Modeling, Sampling, Statistical analysis; Nature of Research—Hybrid Basic-Applied

e) **Deliverables**
VALET software is deployed on PCs at local police stations and available for iPad and iPhone devices.

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Abish Malik, Ross Maciejewski, Timothy F. Collins and David S. Ebert, "Visual Analytics Law Enforcement Toolkit," *IEEE International Conference on Technologies for Homeland Security,* 2010.

**Presentations**

2010 – Timothy Collins, Visual Analytics Law Enforcement Toolkit, IEEE International Conference on Technologies for Homeland Security

January 2011- Ross Maciejewski, Describing Temporal Correlation Spatially in a Visual Analytics Environment, Hawaii International Conference System Sciences

March 2011 – Ahmad Razip, A Mobile Visual Analytics System For Crime Data, DHS University Programs Summit

**Other Reports**

Purdue University feature on undergraduate student Ahmad Razip's work on the iPhone and IPad versions in their "Makers All" campaign

**Research Assistants**
**Current Graduate Students**

SungYe Kim, Purdue University, Electrical and Computer Engineering
Abish Malik, Purdue University, Electrical and Computer Engineering
Sungahn Ko, Purdue University, Electrical and Computer Engineering
Shehzad Afzal, Purdue University, Computer Science
Silvia Oliveros-Torres, Purdue University, Electrical and Computer Engineering

**Undergraduate Students Supported**

Ahmad Razip, Purdue University, Electrical and Computer Engineering
Jing Li, Purdue University, Electrical and Computer Engineering
*Summer Undergrad Research Fellow 2010
Kushal Patel, Purdue University, Electrical and Computer Engineering
*Unpaid through Vertical Integrated Projects course
Shantanu Joshi, Purdue University, Electrical and Computer Engineering
*Unpaid through Vertical Integrated Projects course

**Other Outcomes/Impacts**

This project has shown value for both infield and long-term crime analysis and predictive policing.

**School**
Stanford University

a) **Project Name/Theme**
MDRP5 – Cybersecurity Visual Analytics

Theme 2: Federal Operating Component Projects

b) **PI/Co-PI and Institution**
Prof. Patrick Hanrahan, Stanford University

**Collaborating Partners**
Prof. W. S. Cleveland, Purdue

**Collaborating End-Users**
DHS US-CERT Operations: Lee Rock , Deputy Director(Lee.Rock@dhs.gov)

c) **Research Problem (Abstract)**
We are building tools for visualization of transaction data in the context of (1) internet connection measurements at the packet level arrival timestamps, transport headers, and network headers; and (2) internet connection log files. We seek designs for displays that enable rapid scanning across display pages to enhance the assessment of larger amounts of information than has been possible in the past. In this work we focus on visual methods that enhance both research and operational monitoring. We seek to improve on current practice which relies on reusing data originally designed for billing and network performance purposes.

d) **Technical Approach**
Data Collection—Field monitors; Analytic—Modeling, Statistical Analysis; Nature of Research—Hybrid Basic-Applied

e) **Deliverables**
Isis forensic tool (Java) and documentation delivered electronically to US-CERT

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
This paper appeared in the proceedings of the "12th INFORMS Computing Society Conference (ICS2011), Monterey, CA January 9-11, 2011 pp. 73-90 ISBN 978-0-9843378-1-1

**Research Assistants**
Justin Talbot, Stanford University, Computer Science (grad student)

**Technology Transitions**
ISIS visual forensic tool in pilot deployment for US-CERT Enhanced Analytic Database Capability. ISIS was part of the competitive evaluation for database vendors conducted by US-CERT in Pensacola, FL.
The SSH keystroke detection work has been integrated and shipped in Release 3.0.4 of QoSient's widely-used open source Argus traffic monitoring systems.

**Follow-on Funding/Related Projects**
$88K subcontract through PNNL from DHS S+T , Hanrahan,  9/10-7/11

**Other Outcomes/Impacts**
The two principal accomplishments have been (1) transfer of the Isis forensic tool to a pilot study with the US-CERT Enhanced Analytic Database Capability, and (2) incorporation of our SSH keystroke detection algorithm into Argus Release 3.0.4

**School**
Swansea University

a) **Project Name/Theme**
Uncertainty Visualization in Video Visual Analytics

Theme 6: International Collaborations (also applicable to work on Public Safety Coalition Projects and VASA)

b) **PI/Co-PI and Institution**
Professor Min Chen, Swansea University

**Collaborating Partners**
Professor David Ebert, Purdue University

c) **Research Problem (Abstract)**
In this project, we conducted research into the problem of visualizing multiple video streams in conjunction with a geographical map (which provides an intuitive context). In many situations ranging from traffic monitoring to crowd control, from compound protection to home security), multiple video cameras are used. It is usually difficult to gain an overview of the situation by viewing these independently. We thus need an intuitive and cost-efficient way to view multiple camera streams in a shared context that allows users to relate events in one stream to another at ease.
Our objectives are:
to provide a means to map quasi-3D video information to 2D planar visualization for the integration with geographical maps;
to provide a mechanism that allows portable, rapid and reliable camera-map calibration;
to provide fast visualization algorithms that can present video information in both imagery and abstract visual representations;
to provide a means for estimating uncertainty in video information and to depict it in visualization.

We have made significant advances in this area and thus have achieved these objectives.

d) **Technical Approach**
Data Collection—Field Monitoring; Analytic—Sampling; Nature of Research—Applied

e) **Deliverables**
A prototype system
A small set of algorithms

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Simon Walton, Min Chen and David Ebert, LiveLayer - Live Traffic Projection onto Maps, Eurographics 2011 Posters, Llandudno, Wales, April 2011.

**Other Outcomes/Impacts**

We have developed a real-time method for transforming video information in quasi-3D to spatial information in a 2D planar visualization, which acts as a live layer on top of a conventional satellite image-based mapping system. To facilitate rapid and reliable camera-map calibration, we make use of a semi-automatic calibration scheme. The scheme allows users to pre-define essential projection attributes for each camera in a set-up stage, and map these video streams onto a live layer of the map. We have also developed several schemes for visually mapping information extracted from videos, including traffic speed and density, and uncertainty in illustration. This form of traffic video visualization can potentially reduce the need for specialized traffic sensory devices and infrastructure, and enable better utilization of the existing video-based traffic sensory network.

**School**
University of Houston Downtown

b) **Project Name/Theme**
MSI Collaboration

Theme1: Public Safety Coalition Projects (also has application for work on EE, VASA and IC)

c) **PI/Co-PI and Institution**
Richard Alo, University of Houston Downtown

**Collaborating Partners**
UHD- Richard Alo, Ping Chen, Erin Hodgess, Ongard Sirisaengtaksin
Navajo Technical College- Mark Trebian
California State University Dominguez Hills- Mohsen Beheshti

d) **Research Problem (Abstract)**
The Center for Computational Science and Advanced Distributed Simulation (CCSDS) at UHD, Richard A. Aló, lead, continues MSI outreach to the following: California State University Dominguez Hills, West Houston Center for Science and Engineering of Houston Community College, Navajo Technical College, American Indian Higher Education Consortium, Computing Alliance for HSIs. Also we have assisted in implementation of DHS Scholarship grant for WHC of HCC and their participation in DHS, submission and award implementation of NSF Scholarships program for UHD with recipients working on VACCCINE projects. We provided two minority faculty candidates and students for the DHS faculty /student summer development program- one at VACCINE and other at CCICADA. Efforts were advanced in using a visualization wall with Xbox Kinect technologies to develop control of tele-presence and security access using gesturing and voice commands and sophisticated avatar developments for both NTC and UHD. Such is also under development at CSU-DH (for example: Using visualization walls to simulate crowd behaviors ▪ Mapping/visualizing network attacks from other countries). We are collaborating with Latin American and Caribbean Consortium of Engineering Institutions ([www.laccei.org](www.laccei.org) ) to develop faculty capabilities in Colombia and other South American countries.

e) **Technical Approach**
Data Collection—Data Compiling and Sorting DB, Data Mining, Expert consultation, field monitors, survey, other; Analytic—Case Studies, Modeling, Sampling, Statistical analysis, other; Nature of Research—Applied Basic, Consultation, Coordination/integration, education, Hybrid Basic-Applied, Hybrid Applied—Consultation

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
P. Chen, W. Ding, W. Garcia*, "Adaptive Study Design through Semantic, Association Rule Analysis", (to appear) in International Journal of Software, Science and Computational Intelligence.

P. Chen, A. Barrera*, C. Rhodes*, "Semantic Analysis of Free Text and its Application on Automatically Assigning ICD-9-CM Codes to Patient Records", The 9th IEEE International Conference on Cognitive Informatics, July 7-9, 2010, Beijing, China.

P. Chen, N. Ozoka*, R. Ortiz*, A. Tran*, D. Brown*, "Word Sense Distribution in a Web Corpus," The 9th IEEE International Conference on Cognitive Informatics, July 7-9, 2010, Beijing, China.

P. Chen, W. Garcia*, "Hypothesis Generation and Data Quality Assessment through Association Mining", The 9th IEEE International Conference on Cognitive Informatics, July 7-9, 2010, Beijing, China

Mohsen Beheshti, Nathan Nikotan, Jianchao Han, , *Software Project Scheduling Using A Multi-Agent System,* 8th International Conference on Information Technology: New Generations, ITNG 2011, April 11-13, Las Vegas, NV, USA

**Presentations**
**Student Poster Presentation**
Architecture Design and Implementation of Healthy Heart Application, Nathan Nikotan, Student Research Day, CSUDH Intelligent Network Security Using Genetic Algorithms, Justine Battle, Student Research Day, CSUDH

UHD Student Research Conference 2011, Specialized Data Analysis, Aggregation & Visualization Tool Packages for R", (Hooman Hemmati, Duber Gomez-Fonseca & Sarah Jennisca)

Hooman Hemmati, Duber Gomez-Fonseca and Sarah Jennisca; "Visual Analysis of Large-scale Infornation by Aggregation and Numerical Techniques", presented at TeraGrid 2010, Pittsburg, PA 2010

Hooman Hemmati, Duber Gomez-Fonseca, Sarah Jennisca, Lilian Antunes and Tia Pilaroscia;"Specialized Data Analysis, Aggregation & Visualization Tool Packages for R", CAHSI Annual Meeting, San Juan, PR 2011

**Research Assistants**
Nathan Nikotan, California State University Domiguez Hills
Hooman Hemmati, University of Houston Downtown, Post Baccalaureate

**Number of Undergraduate Students Supported**- 11
UHD: ARP1. Sirisaengtaksin lead ( Karthik Gopavaram, Junior, Maxwell Goedjen, Freshman):
ARP 2. Alo' and Hodgess leads (Hooman Hemmati, Duber Gomez-Fonseca: Lilian Antunres, Sarah Jennisca, Tia Pilaroscia)
ARP3, Chen lead (Walter Garcia, Rafael Ortiz,
ARP 4, Beheshti lead (Nathan Nikotan, Jianchao Han)
ARP 5, Alo, Beheshti, leads with Navajo Tech: Mark Trebian, Jason Arviso, Jared Ribble

**Follow-on Funding/Related Projects**

**Other Outcomes/Results**

**Text visualization:** The goal of this research project is to create an application that allows a user to upload a corpus of text files, then allow the user to search these files for specific words or phrases, and finally display the results graphically. Of particular importance is that graphical information, when presented to the user, remain intuitive with little or no need for legends. Once a search has been performed the user will be presented with a hierarchy of graphical representations:

The highest level will allow the user to quickly understand how often the words and/or phrases are used and in what sender/recipient networks they occur.

The next level will allow the user to inspect a specific network to ascertain which sender/recipient pairs have the highest occurrence of the word/phrase.

The lowest level will allow the user explore all the communications of a specific email address (node) by viewing all the connections to the node.

This application is only relevant when the data corpus is unencrypted text files and assumes that any captured network packets are already decrypted and reassembled.


**Evacuation and Crowd Control:** Crowded buildings become dangerous in emergency situations in which people need to be evacuated quickly. It is important that there be specific evacuation plans in place for every unique building. The motivation for this project is to create simulations to study emergency evacuation situations. So, we can find ways in which crowds can be evacuated safely and efficiently for any given crowded place. Our main objectives of the project are to create a physical model for crowd simulation and a frame work for a 3D graphical simulation tool. In the physical model for crowd simulation, we implement multi agents to represent crowd and fuzzy controllers to control crowd behaviors. We design fuzzy controllers based on behaviors of each agent. Behaviors in our model include panic levels, moving speeds, interaction between agents, reaction to obstacles, and distance between agents. The physical environment/map will be created as a separate entity from the physical. This allows us to replace and modify the physical environment/map efficiently. We expect the model will create two separate output files, agents' locations at different time frames and a physical environment/map.

The 3D graphical simulation tool is designed to accept two files that are generated from the physical model. We anticipated that this design will accommodate any simulations just by changing the 3D models, for example, from a person to a vehicle. Therefore, the 3D graphical simulation tool can be used to simulate traffic of vehicles. Currently, we are working on completing the simulation of physical model. We also started on 3D graphical simulation tool. In this part of the project, we will use XNA to create simulation environment and 3D Studio Max to create 3D models.

In addition to this, we are creating an iOS app to display evacuation routes/exits on each floor inside a building. We developed a tool to show a floor plan on the display of a device such as iPhone. The app will either display a route to the desired destination or display different routes of exits.

**This project is relevant to Theme 4: Event Evacuation.**
Three dimensional (3D) visualization tools have been designed ground up from visual primitives. The basic RGL functionality only allows for the creation of single or compound rectangular surfaces and spheres, and as such necessary primitives for many visualizations such as circles, cylinders and boxes must be manually implemented. The functions for the shapes were designed with arguments that provide extensive modification and customization but simple enough to facilitate collaborative work. These sub tools were programmed by Hooman Hemmati and Sarah Jennisca with practical testing and feedback from Duber Gomez-Fonseca, Lilian Antunes and Tia Pilaroscia.

A number of 3D visualizations were developed and tested. Lilian Antunes and Tia Pilaroscia were instrumental in the design and development of clustered bar-graphs formulated to meet the visualization needs of biological research. The graph allows for the comparison components in multiple control samples from different locations. The visualization allows for a quick overview of component distributions and consistency of samples. A later addition to the visualization was a semi-transparent mesh overlay that would allow for comparison of data summaries of the same or other data sets with current bar graph. Lilian Antunes and Tia Pilaroscia were responsible for sample data, visualization methodology, and designs. Some technical work and programmatic fine tuning was provided by Hooman Hemmati.
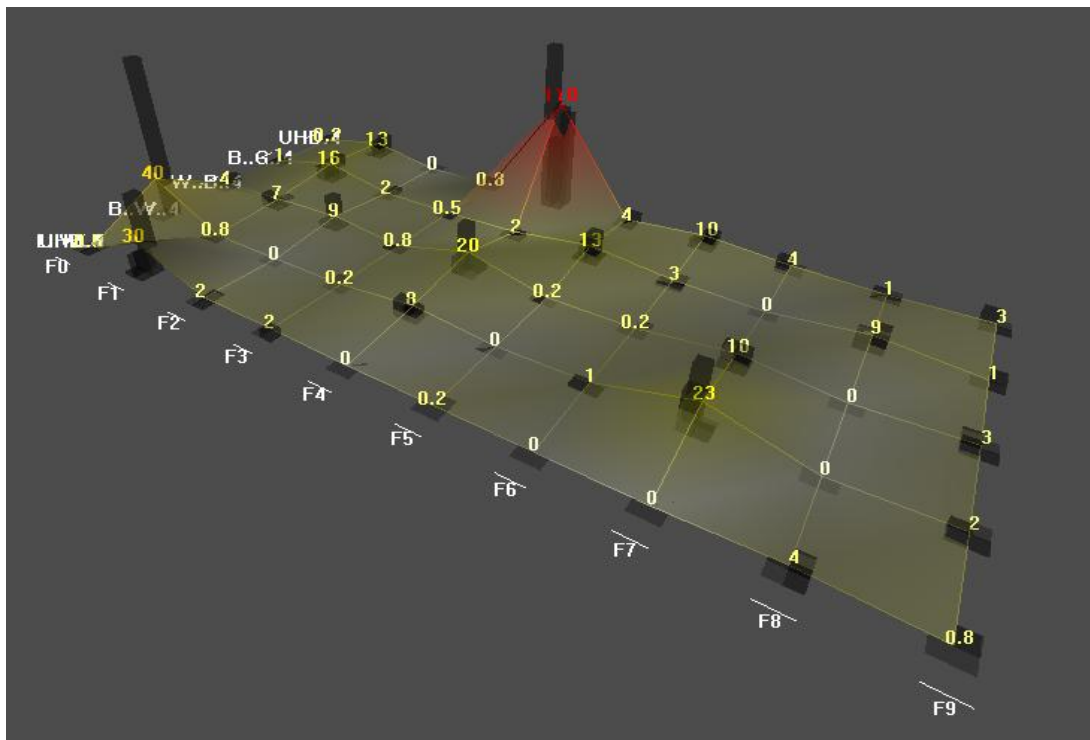


A variety of visualizations are in development and being tested for their application to economics data as well. These include visualization of multiple data-sets as line graphs in 3D space with

transparent green, yellow and red tiers. Text overlay and highlighting is currently in progress. The lines are representing using cylinders to alleviate eye-strain for users, while the transparent tiers indicate alert levels associated with stock behavior. Duber Gomez Fonseca was responsible for sample data, visualization methodology, and designs. Some technical work and programmatic fine tuning was provided by Hooman Hemmati and Sarah Jennisca.

Both visualization can be repurposed for other data sets. For example, the economical data model can easily be used to monitor population growth in various areas or a given time period, or the same area in different time-periods. The bar graph is currently being enhanced so that it can display on a map with each bar in the corresponding geographical location. This model requires the development or adoption of standard source data formatting that includes longitude and latitude data. The model can be used for various purposes, such as modeling levels of viral infections or expenditure patterns by location.

A graphical user interface (GUI) has been designed as an augmentation of the standard Rcmdr interface. Discipline specific drop-down menus are added allowing for the selection of different visualizations. Some options include pop-ups that provide for additional fine tuning by the user. The completed program will also include discipline specific features transparent to user, such as customize aggregation routines optimized for data specific to a particular field. The final product is a point and click research suite that will allow researchers with varying expertise collaborate better by each processing source data using tools designed for their expertise but generating expressive visualization that can communicate across language and disciplinary barriers.

**Tools- Visualization Wall (assisted by California Institute of Telecommunications and Technology- Larry Smarr, Director -** Efforts were advanced in using a visualization wall with Xbox Kinect technologies to develop control of tele-presence and security access using gesturing and voice commands and sophisticated avatar developments for both NTC and UHD. Such is also under development at CSU-DH (for example: Using visualization walls to simulate crowd behaviors ▪ Mapping/visualizing network attacks from other countries). This wall is under construction to be used as a collaborative research /learning environment amongst UHD, CSU-DH and NTC.

University of North Carolina, Charlotte

a) **Project Name/Theme**

Situation Surveillance & In-field Criminal Investigative Analytics

Theme 1: Public Safety Coalition Projects (also applicable to work on FOCP, ERE, and IC)

b) **PI/Co-PI and Institution**

William Ribarsky, University of North Carolina, Charlotte

**Collaborating Partners**

Jianping Fan, KR Subramanian, University of Konstanz, PNNL

c) **Research Problem (Abstract)**

In collaboration with the Department of Justice, PNNL, David Ebert and his group, and several police departments, we have been building new capabilities on a mobile application that is being used for urban routing (taking into account blockages), remote data gathering, communication between emergency responders, and emergency evacuation. A complete GIS-based, client-server system has been developed with a command center interface and multiple mobile devices. A situationally aware evacuation model has also been developed that can be updated based on new obstacles and current conditions. We have developed semi-automated methods to derive routing graphs for large buildings from CAD files, and have now built 3D routing graphs for the entire UNC Charlotte campus (and other buildings with our collaborators). We have developed evaluations of the mobile routing system in action and are now working on multiple testbed exercises with Police partners, nuclear evacuation planners, and PNNL.

We have developed a first version of a visual analytics tool for management, understanding, and action for collections of surveillance videos. As with our other visual analytics methods, this tool is a tight integration of automated methods and interactive displays. It identifies "boundary frames" at the beginning or end of events, provides a level of understanding of events, and permits tracking of moving individuals or objects. We are building towards better automation and better ways of inserting human understanding and direction so that an operator act upon large surveillance video collections. We will share these tools with and get feedback from DHS managers and law enforcement personnel.

d) **Technical Approach**

Data Collection--Compiling & Sorting DB, Field monitors, Survey; Analytic--Case Studies, Modeling, Statistical analysis; Nature of Research--Coordination/integration, Hybrid Basic-Applied, Hybrid Applied –Consultant

e) **Deliverables**

We worked on incorporating our urban routing and emergency response system into the Canada-US Experiment Emergency Resiliency Environment (CAUSE –ERE) scenario set for demonstration at the end of June. Joe Kielman and other DHS managers will be in attendance for the experiment. The initial experiment involves an earthquake in the Pacific Northwest spanning

the U.S.-Canada border. Our system will be used for mobile emergency routing and situational awareness in a downtown Seattle building.

We prepared a shooter exercise for a campus building. In this exercise, a person with a gun takes hostages in a campus building. Ultimately, the person is cornered and barricaded in the building. A campus lockdown results from the incident. This is a re-running of a similar exercise that took place 1.5 years ago in which UNCC Police, Charlotte Police, emergency medical responders, hostage negotiators, and the university executive team took part. One objective of the current exercise is to use the new capabilities of the mobile urban emergency response and evacuation system to address some of the shortcomings uncovered in the previous exercise. Among the issues uncovered were the lack of situational awareness by the executive team (which must make campus-wide decisions) and the lack of a process to manage lifting of the campus-wide lockdown. The UNCC Police will participate in the current exercise and Charlotte Police will be invited as observers.

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Jianfei Liu, Kyle Lyons, Kalpathi Subramanian, William Ribarsky. Semi-Automated Processing and Routing Within Indoor Structures For Emergency Response Applications", Proceedings of SPIE Defense, Security+Sensing, 2010.

**Presentations**
W. Ribarsky, Invited speaker, "Mobile Application for First Response and Emergency Evacuation in Urban Settings." ACM Com.Geo 2010 (Washington, DC, June 2010).

**Research Assistants**
John Merritt, Ph.D., Computer Science, UNCC
Onyewuchi Obirieze, Ph.D., Earth and Geographic Sciences, UNCC
Jack Guest, M.S., Computer Science, UNCC
Praveen Tripathi, Ph.D., Computer Science, UNCC

**School**
University of North Carolina, Charlotte

a) **Project Name/Theme**
MDRP 11: Financial Fraud and Risk Visual Analytics

Theme 2: Federal Operating Component Projects (also applicable to work on ERE and VASA)

b) **PI/Co-PI and Institution**
William Ribarsky, University of North Carolina, Charlotte

**Collaborating Partners**
Bank of America

c) **Research Problem (Abstract)**

In this project, new visual analytics tools are being developed to explore and discover behaviors (suspicious behavior, emerging risk behavior) in financial transaction data. Emerging risk analysis is a broad area that seeks to uncover emerging risk in the bank's transactional, credit card, and loan portfolios that may not be adequately accounted for. The effect of the economic downturn and job losses in specific industries on unsecured loans for specific groups of customers is an example of such emerging risks. To meet this need, we have developed a new tool, RISKVA, in cooperation with risk analysts at Bank of America. To evaluate the efficacy of RiskVA, we conducted expert evaluations with risk analysts from BOA, and found that most analysts considered the tool to be useful and complimentary to their existing analysis needs. We further identified analysis scenarios for which our system could provide analysts with insights to develop appropriate risk management strategies.

Given these results, RiskVA presents substantial qualitative advances over current practices in consumer risk analysis:

It provides a visual exploratory environment to handle consumer risk data that scales to hundreds of thousands of credit transactions over any given length of time.

It provides highly coordinated interactive visualizations to enhance both tactical and strategic risk analyses that are essential in identifying emerging risks.

It provides customizable workspaces that support the individual analyst's analysis routines.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Data Mining, Expert consultation; Analytic—Case Studies, Statistical analysis; Nature of Research--Coordination/integration, Hybrid Basic-Applied

e) **Deliverables**
A procedure for creating simulated bank transaction datasets has been developed. The procedure produces bank transaction records over time that simulates the behavior of real, aggregated

transactions from banks. The datasets can be scaled to any size; anomalous trends or suspicious behavior can be inserted. The procedure is available to generate datasets for DHS partners and customers to test out analysis procedures or to generate data for case studies. Datasets like these are quite hard to obtain.

Our WireVis tool for analyzing bank transactions over extended periods of time is available to DHS partners and customers. We are expanding its capabilities to analyze the unstructured text fields that accompany transactions and that describe the purpose of a transaction, identify payees, and give other information. There are currently no automated methods to study these fields and connect them to the transactional records.

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Presentations**
W. Ribarsky, Invited presentation, "Financial Visual Analytics," INFORMS 2010 (Austin, November, 2010).

**Student Thesis**
Dong Jeong, Ph.D., Computer Science, "Knowledge Visualization: From Theory to Practice"
Remco Chang, Ph.D., Computer Science, "Thinking Interactively with Visualization

**Research Assistants**
Xiaoyu Wang, Ph.D., Computer Science
Matthew Hawkins, M.S., Computer Science

**Follow-Up on Funding/Related Projects**
Bank of America, $90,000, PI: William Ribarsky, "Development of new tools for risk analysis," December, 2009 to May 31, 201

**School**
University of North Carolina, Charlotte

a) **Project Name/Theme**
Multimedia Visual Analytics for Investigative Analysis

Theme 2: Federal Operating Component Projects (also applicable to work with VASA, IC)

b) **PI/Co-PI and Institution**
William Ribarsky, UNC Charlotte

**Collaborating Partners**
Jing Yang, Jianping Fan, University of Konstanz

c) **Research Problem (Abstract)**
We have developed a novel visual analytics framework to support automatic understanding, exploration, and more effective analysis of large-scale online multimedia documents for homeland security applications. Collections of many thousands of online unstructured documents, or streaming accumulation of online content, can reveal themselves and their meanings. Investigators in many areas often use multiple public information sources. More and more, these sources have multimedia content. The analysts are greatly hampered and sometimes stymied by the fact that there are no automated tools for intelligent categorization and analysis of the copious amounts of information that flow from these sources. For example, there are few tools to understand the content of unannotated images. There is hardly anything to extract and relate themes or stories from multimedia content (for example, from broadcast video or multimedia Web pages) and organize them for exploratory analysis. Thus analyzing broadcast news over days or weeks from multiple sources in several languages is untenable because trained analysts would have to actually watch the many thousands of hours of video.

To handle the rich temporal behavior of this online multimedia, we have developed a novel approach that identifies meaningful events as they emerge over time and then structure them into an appropriately clustered collection for exploration and understanding.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Data Mining; Analytic--Case Studies, Modeling, Statistical analysis, Image understanding & Topic Modeling; Nature of Research-- Coordination/integration, Hybrid Basic-Applied

e) **Deliverables**
Our image analytics and exploration tools and source code are released to users at PNNL. They are available to customers in DHS S&T, if desired.

Our automatic object detection and classification tools and source code are released to Vufind.com. They are available to customers in DHS S&T, if desired.

Our multimedia analytics tools including StreamIt, Newdle, and PIWI are available to DHS S&T stakeholders and customers.

Our systems on multimedia analytics have been demonstrated on multiple leading conferences and institutes

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Journal papers**
Luo, J. Yang, M. Krstajic, J. Fan, W. Ribarsky, and D. Keim: EventRiver: An Event-Based Visual Analytics Approach to Exploring Large Text Collections with a Temporal Focus. IEEE Transactions on Visualization and Computer Graphics, to appear.

J. Yang, D. Luo, Y. Liu: Newdle: Interactive Visual Exploration of Large Online News Collections, IEEE Computer Graphics & Application 30(5): 32-41 (2010).

**J. Fan**, Y. Shen, C. Yang, N. Zhou, ``Structured max-margin learning for inter-related classifier training and multi-label image annotation'', *IEEE Trans. on Image Processing*, vol. 20, no. 3, 2011.

H. Luo, **J. Fan**, Y. Zhou, "Multimedia news exploration and retrieval by integrating keywords, relations and visual features", *Multimedia Tools and Applications*, vol.51, pp. 625-648, 2011.

**J. Fan**, **W. Ribarsky**, R. Jain, ``Social computing for collaborative image understanding", *IEEE Multimedia*, vol. 18, 2011 (invitation submission, to appear).

Y. Shen, N. Zhou, **J. Fan**, ``Integrating bilingual searches for junk image filtering", *Multimedia Tools and Applications*, vol.52, 2012 (invitation submission, to appear).

Y. Guo, X. Lin, Z. Teng, X. Xue, **J. Fan**, ``A covariance matrix-free iterative algorithm to achieve distributed principal component analysis on vertically partitioned data", **Pattern Recognition**, vol. 44, no. 11, 2011 (to appear).

Y. Zhang, C. Yang, Y. Shen, N. Zhou, J. Fan, H. Luo, "A cross model approach to cleansing weakly tagged images", *IEEE Multimedia*, vol.17, no.4, pp.18-25, 2010.

Y. Shen, J. Fan, "Multi-task multi-label multiple instance learning", *Journal of Zhejiang University SCIENCE C* (Computers & Electronics), ISSN 1869-1951, ISSN 1869-196X, 2010 (invitation paper).

J. Peng, N. Babaguchi, H. Luo, Y. Gao, **J. Fan**, "Constructing distributed hippocratic video databases for privacy-preserving online patient training and counseling", **IEEE Trans. On Information** *Technology in Biomedicine*, vol.14, no.8, 2010.

Z. Li, **J. Fan**, "Exploit camera metadata for enhancing interesting region detection and photo retrieval", *Multimedia Tools and Applications*, vol. 46, no.2-3, pp.207-233, 2010.

Q. Luo, X. Kong, G. Zeng, **J. Fan**, ``Human action detection via boosted local motion histograms'', ***Machine Vision Applications***, vol.21, no.3, pp.377-389, 2010.

H. Luo, **J. Fan**, ``Multimedia news exploration and retrieval by integrating keywords, relations and visual features'', ***Multimedia Tools and Applications***, vol.46, no.11-12,

**Conference papers**

W. Zhang, X. Xue, J. Fan, ``Correlative multi-label multi-instance image annotation", **IEEE ICCV** (IEEE Intl Conf. on Computer Vision), Barcelona, 2011.

N. Zhou, Y. Shen, J. Peng, X. Feng, **J. Fan**, ``Leveraging auxiliary text terms for automatic image annotation'',  ***ACM WWW*** (Intl World Wide Web Conferences), Hyderabad, India, 2011.

W. Zhang, X. Xue, **J. Fan**, ``Multi-kernel multi-label learning with max-margin concept network'', **IJCAI** (22th Intl. Joint Conf. on Artificial Intelligence), Barcelona, 2011.

Y. Zhang, C. Jin, L. Cen, X. Xue, J. Fan, "Fusion of multiple features and supervised learning for Chinese OOV term detection and POS guessing", ***IJCAI*** (22th Intl. Joint Conf. on Artificial Intelligence), Barcelona, 2011.

Y. Zhang, C. Jin, L. Cen, X. Xue, **J. Fan**, ``Learning inter-related statistical query translation models for English-Chinese bi-directional CLIR", ***IJCAI*** (22th Intl. Joint Conf. on Artificial Intelligence), Barcelona, 2011.

Y. Nakashima, N. Babaguchi, **J. Fan**, "Automatic generation of privacy-protected videos using background estimation", ***IEEE ICME*** (IEEE Intl. Conf. on Multimedia and Expo), Barcelona, 2011.

C. Yang, X. Feng, J. Peng, **J. Fan**, "Efficient large-scale image data set exploration: Visual concept network and image summarization", ***MMM'11*** (Intl. Conf. on Multimedia Modeling), pp. 111-121, 2011.

N. Zhou, J. Peng, X. Feng, **J. Fan**, "Towards more precise social image-tag alignment", **MMM'11**, (Intl. Conf. on Multimedia Modeling), pp. 46-56, 2011.

Y. Shen, **J. Fan**, "Leveraging loosely-tagged images and inter-object correlations for tag recommendation", **ACM Multimedia**, 2010.

Y. Nakashima, N. Babaguchi, **J. Fan**, "Automatically protecting privacy in consumer generated videos using intended human object detector",  ***ACM Multimedia***, 2010.

**J. Fan**, Y. Shen, N. Zhu, Y. Gao, ``Leveraging large-scale weakly-tagged images from Internet'', **IEEE CVPR** (IEEE Conf. on Computer Vision and Pattern Recognition), 2010.

X. Xue, H. Luo, J. Fan, "Structured learning for automatic image annotation",  ***ACM CIVR'10 (***ACM Conf. on Image and Video Retrieval), 2010.

C. Yang, J. Fan, "Speed up duplicate/near-duplicate image detection", ***ICIMCS'10*** (The Second International Conference on Internet Multimedia Computing and Service), 2010.

Y. Shen, N. Zhou, **J. Fan**, "Integrating bilingual searches for junk image filtering", ***ICIMCS'10*** (The Second International Conference on Internet Multimedia Computing and Service), 2010.

N. Zhou, Y. Shen, **J. Fan**, ``Automatic image annotation by using relevant keywords extracted from auxiliary text documents", ***ACM Multimedia Workshop on Large-Scale Image Retrieval,*** 2010.

J. Alsakran, Y. Chen, Y. Zhao, J. Yang, and D. Luo: STREAMIT: dynamic visualization and interactive exploration of text streams. Proc. IEEE Pacific Visualization Symposium 2011, pages 131-138.

J. Alsakran, Y. Zhao, D. Luo, and J. Yang: Visual Analysis of Stream Texts with Keyword Significance. Poster of InfoVis 2010.

Y. Chen, J. Alsakran, S. Barlowe, J. Yang, and Y. Zhao: Supporting Effective Common Ground Construction in Asynchronous Collaborative Visual Analytics. Proc IEEE Conference on Visual Analytics Science and Technology 2011, conditionally accepted.

**Presentations**

W. Ribarsky, Invited Speaker, "A Mixed-Initiative Visual Analytics Approach to Topic Modeling in Research Collections," AAAS-NSF Workshop on Science Policy (October, 2010).

Ribarsky, W. Distinguished Lecturer, "Developing a Visual Analytics Approach to Analytic Problem-Solving, " School of Computational Science and Engineering, Georgia Tech, February, 2010.

J. Fan. WOCC'10: The 19th Annual Wireless and Optical Communications Conference (WOCC'10), Shanghai, China, 2010. Speech Title: A visual analytics framework for supporting human-centered multimedia computing

J. Fan. ICMR'10: International Workshop on Multimedia Retrieval (ICMR'10), Amsterdam, 2010. Speech Title: Interactive exploration of large-scale image/video collections

J. Fan. SMLCMR'11: Intl. Conf. on Statistical Machine Learning and Cross-Media Retrieval (SMLCMR'11), Hangzhou, 2011. Speech Title: Cross-modal cleaning and exploration of large-scale social images.

**Research Assistants**

Chunlei Yang, Ph.D., Computer Science
Yi Shen, Ph.D., Computer Science
Yang Chen, Ph.D. Computer Science
Yujie Liu, Ph.D., Computer Science
Dongning Liu, Ph.D., Computer Science
Scott Barlowe, Ph.D., Computer Science

University of North Carolina, Charlotte

a) **Project Name/Theme**

SP 14-Task 5, SP 15-Task 2: Visual Analytics for Security Applications (VASA);

Theme 5: Visual Analytics for Security Applications (also applicable for work in IC)

b) **PI/Co-PI and Institution**

William Ribarsky, UNC, Charlotte

**Collaborating Partners**

Bill Tolone, University of Konstanz
University of Stuttgart
Purdue University

**Collaborating End-Users**

Duke Energy
Siemens

c) **Research Problem (Abstract)**

During this period, this project has been in planning phase and full funding has not been allocated yet. Our research, which focuses on the use of interactive visualizations to explore complex, high dimensional, large volume data as a means of empowering human cognition and facilitate analytical reasoning, provides a foundation for the proposed efforts. The proposed effort will provide an approach to critical infrastructure and cascading effect modeling (e.g., due to key outages or overloads) fully integrated with interactive visualization to produce a true visual analytics decision-making environment. Among the capabilities of this visual analytics environment will be the ability to identify and understand developing cascading effects quickly in terms of their temporal and multidimensional behavior and their principal relationships (including cause and effect). The visual analytics approach will identify key events in sequence. One aspect of our research will be to turn these identified events, their behaviors, and relationships into actionable plans or courses of action, complete with annotations provided by the analysts and augmented with alternative scenarios as an emergency unfolds. Project team members have worked closely with U.S. Government analysts to understand requirements and collect feedback. Research results have led to deployable desktop and net-centric solutions for the U.S. Government. In this project, we are planning to augment work with the U.S. Government by working with Duke Energy. We plan to develop testbeds in North Carolina, Ohio, and Indiana in conjunction with Duke Energy and our partners at Purdue. These efforts will be coordinated with the German efforts.

Initial infrastructure modeling for the North Carolina Charlotte and coastal regions has begun.

d) **Technical Approach**

Data Collection--Compiling & Sorting DB, Data Mining, Expert consultation, Field monitors, Survey; Analytic--Case Studies, Modeling, Sampling, Statistical analysis; Nature of Research—

Applied, Basic, Consultation, Coordination/integration, Education, Hybrid Basic-Applied, Hybrid Applied –Consultation

e) **Deliverables**
A complete scenario for a large scale coastal event (hurricane) with attendant storm surge, flooding, and extreme rainfall moving inland.

A visual analytics system to explore the character, relationships, and results of the critical infrastructure model for this or other scenarios. This system will be available to DHS S&T stakeholders and customers.

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Presentations**
W. Ribarsky, Invited Participant, U.S. - German Workshop on Visual Analytics for Security Applications (Berlin, February, 2010)

**Research Assistants**
Khaldoon Dhou, Ph.D. student, UNCC
Osarieme Omokaro, Ph.D. student, UNCC

University of North Carolina, Charlotte

a) **Project Name/Theme**
SP 15-Task 3: Event Evacuation Planning

Theme 1:  Public Safety Coalition Projects (also applicable to work with ERE, EE)

b) **PI/Co-PI and Institution**
William Ribarsky, UNC, Charlotte

**Collaborating Partners**
KR Subramanian

**Collaborating End-Users**
UNC Charlotte Police
Charlotte Mecklenburg Police

c) **Research Problem (Abstract)**
In collaboration with the Department of Justice, and several police departments, we have created a situationally-aware evacuation modeling and decision-making capability. This capability is developed in conjunction with the mobile 3D routing client-server system described in our report for MDRP 3. The evacuation model gives optimal evacuation routes based on number and distribution of building occupants. It can be applied to multiple urban buildings and the surrounding street network. The evacuation plan can be updated based on new events such as blocked exits or unexpected evacuee traffic. Since the thousands of times steps produced in a typical model run will be of limited immediate use to emergency responders, we are developing a method to automatically extract the key events in the evacuation model results (e.g., where and when key points of congestion occur around stairwells and exits) so that they can clearly see where to deploy responders and what to do.

Through the command post of the client/server system multiple first responders can send location, search, images, evacuation status, and other information. Thus the system provides a detailed, comprehensive view of a large evacuation as it unfolds. Our goal is to show the constantly evolving, large scale first response and emergency evacuation scenario in an understandable way with key events identifiable. We are developing models for large arenas, conference facilities, and stadia that will be embedded in urban models.

d) **Deliverables**
Situationally-aware evacuation model with several building examples and ability to expand to new buildings. The model is available for use by DHS S&T stakeholders, customers, and to partners such as PNNL.

We have begun to prepare for an evacuation exercise built around a scenario of a nuclear emergency at the McGuire Nuclear Power station, about 15 miles from the UNC Charlotte campus. Emergency planners and responders, including UNCC Police, Charlotte Mecklenburg Emergency Management, and FEMA observers, will use the on-campus field house as an

evacuation site. The exercise will take place in Summer, 2011, and we will provide a report of results to DHS.

e) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Presentations**
W. Ribarsky, Invited speaker, "Mobile Application for First Response and Emergency Evacuation in Urban Settings." ACM Com.Geo 2010 (Washington, DC, June 2010

**Research Assistants**
Jack Guest, M.S. student, UNCC
Seth Clark, M.S. student, UNCC

**School**
University of Washington

a) **Project Name/Theme**
MDRP 7: Introducing Sustainable Visual Analytics into Command Center Environments

Theme 2: Federal Operating Component Projects (also applicable to work with VASA, IC)

b) **PI/Co-PI and Institution**
Mark Haselkorn, University of Washington

**Collaborating Partners**
Texas A&M, TrueStone Inc., Health & Human Services, World Vision International, Simon Fraser University

**Collaborating End-Users**
Area Maritime Security Committer (Research Leadership Council), USCG Sector Puget Sound, Customs and Border Protection, U.S. Army 833rd Transportation Battalion, Port Authorities for Tacoma, Seattle, Olympia

c) **Research Problem (Abstract)**
First, in partnership with a small business (Truestone, Inc.), Texas A&M and numerous DHS client agencies, we advanced an integrated precision information environment for use within the critical decision-making processes of a multi-agency regional security community. A simulated "worst-case" table-top exercise was conducted using "Truevue" to present a degrading scenario within a common operational environment among the participants.

Second, in partnership with the National Center for Border Security and Immigration, we competed, presented and published a three-site study of border security operations and the role of command, control and communication (C3) systems in support of those operations. We outlined the competing tensions inherently intertwined with C3 systems in the complex, critical workflow and decision-making processes within the border security mission. C3 systems not consciously aligned with desired practices and decision-making can have negative implications for CBP workflow and decision making.

Third, in partnership with World Vision International, we explored how logistics stakeholders in a large international humanitarian organization experience and perceive speed of relief and emergency response operations. We found the experience of speed is often comparative, not solely objective; close communication between internal clients (field requestors) and service providers (logistics team) can make clients more likely to experience the logistics process as fast; and feeling in control of decision-making can make both clients and service providers more likely to experience the logistics process as fast.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Expert consultation, Field monitors, Simulation; Analytic—Case Studies, Sampling, Statistical Analysis; Nature of Research—Applied, Basic, Consultation, Coordination/integration, Hybrid Applied—Consultation

e) **Deliverables**

| Milestones for VACCINE Period ? | 06/01/10 – 05/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| Advance the visual analytic tool developed last year to be used for the use of advancing knowledge and facilitating complex decision-making in a multi-jurisdictional environment | X | X | | |
| Develop and conduct a regional port security scenario with stakeholders for at table- top exercise simulation | | X | | |
| Observe and analyze the use, benefits and system in diverse command and control environments | | X | X | X |
| Conduct initial usability studies of Truevue | | X | | |
| Explore and pursue Phase II DHS support for the commercialization of the Truevue system | | X | X | |
| Incorporate lessons learned and usability study findings into VA development | | | X | X |
| Begin next steps with Port Research Committee for use of Truevue in ongoing exercise | | | | X |

f) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Walton, R., Mays, R., & Haselkorn, M. (2011, May). Defining fast: factors affecting the experience of speed in humanitarian logistics. *Proceedings of the 2011 international conference on international systems in crisis response and management*. Winner of 2011 ISCRAM Best Paper Award

Mays, R., Zachary, M., Murat, A., Haselkorn, M. (2011, May). Aligning Border Security Workflow and Decision Making with Supporting Information and Communication Systems, *Proceedings of the 2011 international conference on international systems in crisis response and management*

 *Note: Proceedings of the 2011 International Conference on Information Systems for Crisis Response and Management (ISCRAM2011)* are fully-refereed, blind, 45% acceptance.

**Presentations**
*Military Outload Exercise After Action Review,* Puget Sound Port Readiness Committee, October 14, 2010, presented by Prof. Mark Haselkorn.

Aligning Border Security Workflow and Decision Making with Supporting Information and Communication Systems, 2011 ISCRAM Conference – Lisbon, Portugal, May 9, 2011, presented by Prof. Mark Haselkorn.

Defining "Fast": Factors Affecting the Experience of Speed in Humanitarian Logistics, 2011 ISCRAM Conference – Lisbon, Portugal, May 11, 2011, presented by Robin Mays.

**Other Reports**
Usability Report on Truevue Precision Information Environment.

Final Report Phase I/Proposal Phase II for Research on Precision Information Environments.

Research Leadership Council White Paper on Intermodal Impact to Recovery.

**Student Thesis**
Brian LeBlanc, Information School, Masters Senior Project:
*Information Sharing in the Seattle Emergency Operations Center*

**Research Assistants**
Robin Mays, University of Washington, HCDE Masters Program (recently accepted into Ph.D. program)
Alena Benson, University of Washington, HCDE M.S. degree, June 2011

**Follow-on Funding/Related Projects**
Principal Investigator Peter Marsh, DHS small business grant # DHS SBIR-2010.1-003, $100,000, (Mark Haselkorn PI on sub-award of $16,666 to UW), May – December 2010.

Principal Investigator Mark Haselkorn, UW component of the "National Center for Cognitive Informatics and Decision Making in Healthcare (NCCD), Office of the National Coordinator for Health Information Technology,  $1,100,000 over four years (of $15,000,000 four-year award to the University of Texas at Houston Medical Center), April 2010 – March 2014.

**Other Outcomes/Impacts**
*In this section, we will focus on the results and accomplishments made in the advancement work done around the development of an integrated dashboard framework.  Under an SBIR led by Truestone, Inc., we teamed with FAZD (Texas A&M) and VACCINE (Purdue) to deliver and assess the second iteration of a next-generation CCI system. Thanks to our ongoing collaboration with the Puget Sound Port Readiness Committee (PRC), we were able to plan and conduct a computer supported "worst-case" table-top exercise in the Fall using a "Truevue dashboard" to present the scenario and share information among the participants.  There was general agreement within the PRC that there would be significant benefit from such an exercise, especially considering that there would be a change of command within the Army's 833rd Transportation Battalion during the summer.*

**Baseline Demonstration – The Table-top Exercise**
As part of the Phase I effort, an initial baseline demonstration tool (i.e., Truevue) was developed using Texas A&M's Information Dashboard Framework (IDF). This tool served as the basis of the regional table-top exercise that was conducted. Truevue provided the participants involved with a common operational environment and allowed for sharing of information across different agencies and echelons. In addition, all injects delivered within the exercise were done using this system. These injects included messages providing details about the situation as it unfolded, documents and images describing the incident, streaming video from the field, newsfeeds

describing major events both past and present, and geospatial information identifying key locations of interest. This system was backed by a separate simulation and exercise management tool developed by Texas A&M that allowed for design and execution of the underlying scenario, and simulation of all resources and entities involved with the incident. The Truevue system provided the participants a virtual portal into the incident taking place, and supported both situational awareness and decision-making by the participants.

**Planning**

In the Oct 14, 2010 Table-top Exercise with the Puget Sound Port Readiness Committee, University of Washington (UW) researchers in partnership with Truestone and TAMU intended to determine how the precision information environment (PIE) developed at the Texas (TEES) influenced the transition from independent organizations with autonomous authorities to a unified command in response to a situation that crossed jurisdictional boundaries. Did PIE aid in the recognition of this need? Did PIE help to facilitate and simplify that transition?

**Scenario Development**

The exercise team consisted of representatives from UW, TAMU, Truestone, the 833rd Transportation Battalion, the USCG Sector Puget Sound and the Port of Tacoma (APM Terminals.)  Truestone and the UW team led the exercise development, conducting numerous in-person interviews and follow-up confirmatory discussions with subject matter experts and operational planners from the participating organizations in order to understand their particular mission processes, ways of communicating and information flow, authorities, and relationships. This interview and confirmation sequence was used to ensure realistic scenarios and stimuli injects that mimicked the usual way of operating and that individuals and organizations would obtain information via the expected method (phone, email, official message, news, specific C2 system, or by chance.)  We conducted preliminary interviews with each of the participating organizations to adjust the scenario to meet reality.  For example, during the interview with the second Military Sealift Command we discussed our intention within the scenario to pick up cargo in San Diego for off-load in the Puget Sound.  The ship we had identified as the second participating vessel, however, was built outside the United States, so provisions of the Jones Act would have prohibited it from moving cargo between two US ports.  While a scenario inject such as this would not have affected exercise play, it would have affected our credibility.  Scenarios were predominantly constructed to flow in a way that participants would have had experience. Our going in position was that there would not be any catastrophic events such as an earth-quake, nor any tricks to fool the participants.

The planning team was able to successfully prompt the participants in creating a shared information environment in two key ways:
There existed of a large amount of information and the resulting need for participants to filter, select and determine whether to access, share, or act.

The information came to participants in the way they expected to receive it, and included certain critical pieces of information owned by one key member of the community upon whom the rest of the community would rely on to share that information.  For example early in the exercise play, the representative from the Navy's Military Sealift Command received a phone call indicating that the second vessel assigned to the Military Outload was experiencing mechanical trouble while en route to the Puget Sound.

The scenarios we developed were designed to create an operational environment that would lead to dynamic reorganization among the participants that we hoped to observe – both formal as in establishing a Unified Command and informal as in determining appropriate courses of action in response to a delay in the Military Sealift Command's second vessel's expected arrival or an unfortunate incident with the notional protesters overturning in their kayaks, requiring a Search and Rescue. (The value of the "storyline" was demonstrated not only in the TTX, but also in that it was adopted by a large consortium for part of a subsequent proposal to IARPA.)

The primary scenario "storylines" we developed and the corresponding dynamics we sought to inspire included:

| Scenario | Desired Dynamic Event | Goal Result Achieved |
|---|---|---|
| **Credible Threat of Terrorist Action** | Formation of Unified Command | Y |
| **Man Overboard** | Conflict for USCG security vs. safety | Y |
| **Ramp Accident** | Army, Port Collaboration | Y |
| **Security Breach** | Transition of Authority | Y |
| **First Lady Visit** | Resource Limitations/Request for Local Assets | Y |
| **Ship Delay** | Collaboration for Contingency Plans; notifications to higher headquarters | Y |

**Table 1: Results of Scenario Events**

**Exercise Play**

The table-top exercise was conducted at the University of Washington's Human Centered Design and Engineering computer lab. Participants were organized physically near those they would be collocated with and some physical separation from those that would not be collocated. The exercise simulated three days of a military outload (MOL) operation at the Port of Tacoma. Appendix II provides the advance exercise brochure distributed to the participants. Appendix III presents the reference sheets given to all the participants and role players at the beginning of the exercise to ensure a common understanding of the characteristics of the Dashboard framework. Appendix VI provides the coordinating instructions and situation manual. There was also a pre-exercise briefing.

The actual exercise consisted of about 30 participants and 15 role players. Role Players were played by subject matter experts, many of whom were consulted in the creation of the scenarios. It is worthy to note that to provide a realistic event for participants, it was key to have experienced experts as role players who had the ability to realistically react to participants natural responses to the injects. This also was critical to keeping the exercise moving forward naturally and at a realistic pace.

The scenario scripts were injected directly into the PIE framework via the Texas A&M applied science team. All communications were directed to occur in chat in order to capture the coordination and sequence of the event for later analysis. Timing was monitored by the exercise lead and could be in step with where the group had advanced. The exercise lead also served as

'RADAR' a helper they could access via chat or email for when participants found themselves 'stuck' due to artificialities within the exercise.

The exercise consisted of a morning briefing on the operational details of the ongoing MOL. This was followed in the lab by a one hour window of orientation to features of PIE and a few minutes to get situated to their workstations. The exercise then consisted of 90 min progressions of an exercise day followed by a 15-20 minute debrief time for reflection built in. Discussion time was facilitated over the long lunch time.

In addition, Truevue allowed for custom sharing of information between different users and groups. A user had the ability to determine what information to share and when. An example case of this took place during the exercise when the Coast Guard released the information about the location and number of security vessels it had deployed to support the mission. Prior to this release, this information was not known by all agencies. Truevue allowed the users to see both the details and specific locations of these vessels once the information was shared.

**Transition to Unified Command**

By far one of the most beneficial elements of the exercise was the ability for the community to consider and reflect on the challenges and implications of the transition to the unified command. It naturally led to a discussion of how the PIE was useful and could be more useful in this transition. The discussion around 'who's in charge' precipitated almost immediately within the mid-day group discussion. One participant said "this is a military outload (MOL), DOD is the customer..hence, DOD is the Supported Commander." The MOL commanding officer was quick to clarify he would not be in charge of anything outside of the outload. The participant differed, drawing out the order of command on the board. The ranking USCG sector representative outlined the four conditions to act as unified commander, clarifying that the role could only fall to the USCG Sector. Another participant pointed out that actually the Port Authority could qualify to be unified commander.

This discussion is exactly the type of tension we had hoped to elicit. It came earlier than expected, however when the participants returned to the exercise the discussion set the stage for an incredibly friendly and smooth collaboration during the transition to accept and align to the new authoritative roles. Further consideration of how the system could aid in the transitions between authorities were raised, to include questions around how the technology could help the community to anticipate unified command and facilitate the transition in advance of crisis.

**Lessons Learned**

Through the use of the Truevue system during the tabletop exercise, a great deal of valuable information was obtained to help guide the future work performed. The general consensus across the exercise participants was that this would be a very beneficial tool to allow them to share information and offered many different capabilities that do not currently exist today. Many of these capabilities would help with both their day-to-day operations, and their response capabilities in the event of a major incident. It was evident that several areas of future work would be necessary including proper scalability of the system to allow for multiple 'owned' installations, increased security for controlled management of information, and capabilities to allow for different agency administration and customization of installed systems.

*The Importance of group discussion time*
The points of pause for reflection were extremely valuable and should be replicated. These sessions provided opportunity for participants to provide feedback, share differing perspectives, express frustrations and challenges, and come to better understand their colleagues flexibilities and limitations, and the overall bigger picture. For our purposes, we were able to gain a richer understanding of the sense-making of individuals and the greater group.

*More Robust Training*
The decision was made to provide a one-hour orientation to the system just prior to the exercise in respect for the high demand of the participants' time. While this allowed us to observe things that take a bit to learn intuitively, and contributed to the user's feeling of crisis, it was clear a more thorough, perhaps half-day training would have been more sufficient to more effectively exercise the collaborative and decision-making value of PIE.

*Advance Chat Feature*
While, the participants were encouraged by the opportunity to connect to one another in this way, there was a lot of frustration expressed regarding the chat function. Out of the earlier testing of PIE with the community, there was an overwhelming request to have a chat function to communicate easier. Chat was therefore added for this iteration of use. It should be noted, hence, that this was the first time the chat function was being used outside of development and as the central tool for all communication within this exercise. This frustration was echoed in the user testing as well. While the chat function was not specifically tested, when users were asked what was most frustrating about the test, users unanimously returned to mention the chat function from the exercise. The complaints provided were not around anything considered difficult fixes, but unfortunately it affected the experience, and perhaps took away from providing quality feedback about other features within the PIE. We artificially forced the participants to use the chat so we would have a record of the communications, but ordinarily they would use many different communication methods. The chat is, in any case, a valuable tool for an information sharing framework.

## Conclusions
*Positive impact on decision-making*
The exercise and usability testing revealed an almost unanimous view of the framework improving the ability to make decisions in a complex, multi-jurisdictional, crisis environment.

*Cross-cultural/attitude breakthroughs*
The exercise, itself, but especially through the regular discussions held at the close of each simulation day, led to an increased cross-cultural understanding of different organizations. We believe making the space (be it physical or technological) for various stakeholders within security and disaster response communities cannot be over-valued for its impact on improved understanding and cross-cultural awareness.

*Hubs of Communication*
As indicated by the high spikes on the graph below, certain people in the exercise became communication "hubs" with exceedingly high rates of messages being received and/or sent. The nature of these "hubs" – what makes them occur and the impact on individuals and the group – warrants additional study.

# MOLs are communication intensive
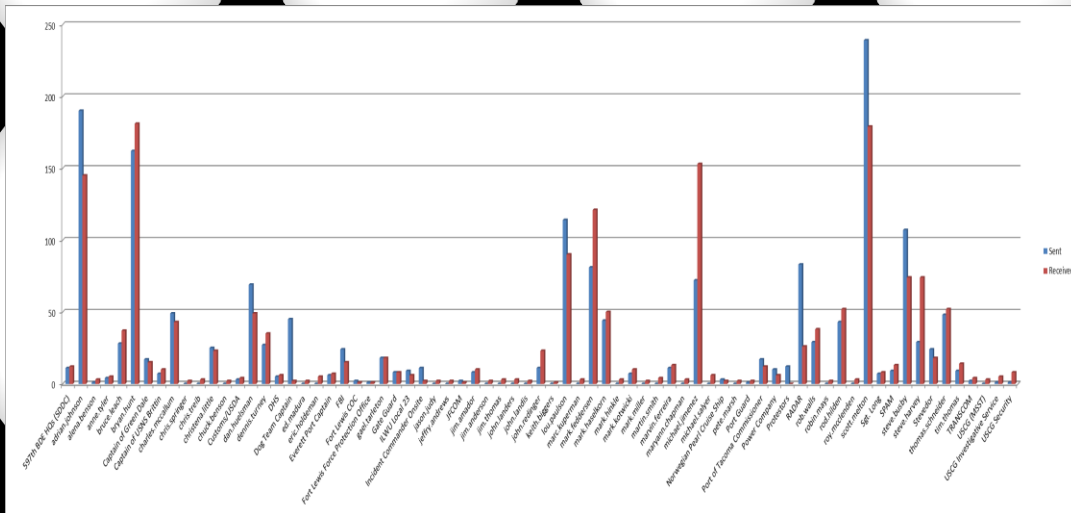


- As situation becomes less "normal," interagency communication increases.
- Certain people become communication "hubs."
- Information overload is real but not for all

*Sustainability*

The PIE offers a unique way of **not** replacing other systems, but sharing them—actually contributing to sustainability of other existing C2 systems. This combined with a user-centered and iterative approach to development translates to greater hope of sustainability. A final consideration that was raised in the discussions was "who is going to own the system." Traditionally, this is critical for the continued adaptation of a system with a growing, advancing organization or community, and we believe it will be for PIE's sustainability as well.

*Commercialization*

Perhaps the most challenging issue in the transition from Phase I to Phase II was the commercialization strategy. Without a commercialization plan, it would be difficult to pursue an overarching governmental Dashboard that would support contingency situations such as the oil spill in the Gulf of Mexico. However, the shared nature of the system makes ownership problematic. In the absence of government ownership, there are a number of regional entities, such as the Port Readiness Committee, Area Maritime Security Committee, or the Pacific Northwest Economic Region, that could sponsor a dashboard capability like Truevue. While many organizations have their own internal IT solutions for information and data sharing, few organizations operate solo in today's increasingly interdependent environment. One avenue that should be further evaluated is a subscription service that provides a range of visual analytic tools in specific situations – one example might be the use of a plume model that assists a decision maker in reviewing options in a response. Applications such as this could reside in a library or catalogue and be invoked when needed. In this way, scarce development funds would not have

to be expended solving the same problems for individual organizations, but could be shared commonly across a range of users.

**School**
The University of Texas at Austin

a) **Project Name/Theme**
Parallelizing Video Algorithms for Real-Time/Near-Real-Time Public Safety Applications

Theme 1: Public Safety Coalition Project (also applicable to work in EE and VASA)

b) **PI/Co-PI and Institution**
Kelly Gaither, Texas Advanced Computing Center, The University of Texas at Austin

**Collaborating Partners**
Dr. Ed Delp, Purdue University

c) **Research Problem (Abstract)**
The proposed effort is to take Dr. Ed Delp's video analytics algorithm and parallelize it on a large-scale high-performance-computing cluster. Parallelizing the video analytics algorithm is needed to try to achieve real-time or near-real-time performance, thus making it useful in those public safety applications that need a real-time response. The outcome of this project will be a parallelized version of Delp's video analytics algorithm and will explore parallelizing on both traditional CPUs and GPUs.

d) **Technical Approach**
Data Collection--Compiling & Sorting DB, Data Mining, Expert consultation, Field monitors, Survey, Other; Analytic—Case Studies, Modeling, Sampling, Statistical analysis, Other; Nature of Research—Applied, Basic, Consultation, Coordination/integration, Education, Hybrid Basic-Applied, Hybrid Applied –Consultation

e) **Deliverables**
The project is just getting underway. The milestones for the project are:
1. Obtain video analytics algorithm from Dr. Ed Delp.
2. Port the serial code to Longhorn and complete performance numbers.
3. Identify areas of the code that can be parallelized.
4. Parallelize each viable unit and document performance numbers/improvements.
5. Demonstrate and deliver parallelized video analytics algorithm and present performance improvements.

I met with Dr. Ed Delp and his students in the Visual Surveillance group when I traveled to Purdue in March 2010. At that meeting we discussed the algorithms the group uses and they provided publications that explained the algorithms in detail. We are currently implementing the video surveillance/feature detection schemes for video on Longhorn in parallel in an effort to run the algorithm and get a real-time response.

**School**
Virginia Tech

a) **Project Name/Theme**
Co-Located Collaborative Analysis on Large, High-Resolution Displays Using Multiple Input Devices

Theme 1: Public Safety Coalition Projects

b) **PI/Co-PI and Institution**
Chris North (Virginia Tech)

**Collaborating Partners**
Lauren Bradel (Virginia Tech)
John Stasko (Georgia Tech)

**Collaborating End-Users**
Local law enforcement (West Lafayette Police Department, Lafayette Police Department)

c) **Research Problem (Abstract)**
We set out to study how co-located collaborative analysis is conducted on large, high-resolution displays in order to evaluate how current visualization tools can be used by outside institutions in a collaborative manner. We conducted a user study to evaluate Jigsaw, a visualization tool developed at Georgia Tech, benchmarked against analytical tools. The study set-up included a large, high-resolution display (108.5 in. x 35 in.; 10,240 x 3200 pixels) with two users seated in front of it with their own mouse and keyboard that could interact independently and simultaneously with the display.

We collected data from the study through solution reports, semi-structured interviews, mouse-log data, and screenshots taken in 15-second intervals. After analyzing this data, we found that analytical tools greatly affect collaboration styles as well as how they interact with the space. We also investigated the display space usage, user roles, territoriality, and the effectiveness of multiple input devices. Using these findings, we made recommendations for improvement to the Jigsaw team.

d) **Technical Approach**
Data Collection--Solution reports, Semi-structured interviews, Mouse log data, Screenshots, Audio/video recordings; Analytic—Case Studies, Statistical analysis; Nature of Research—Basic

e) **Project Outcomes (Results, Accomplishments, Conclusions)**
**Publications**
Vogt, K., Bradel, L., Andrews, C., North, C., Endert, A., Hutchings, D. "Co-located Collaborative Sensemaking on a Large High-Resolution Display with Multiple Input Devices." Conference on Human-Computer Interaction (INTERACT '11). Lisbon, Portugal, September 2011, 17 pages.

**Presentations**
July 2010, Katherine Vogt, "Co-located Collaboration on a Large, High-Resolution Display," REU Symposium.

September 2011, Lauren Bradel, "Co-located Collaborative Sensemaking on a Large High-Resolution Display with Multiple Input Devices," INTERACT '11 (Paper presentation)

November 2011, Lauren Bradel, "Benefits of Co-located Collaboration on Large, High-Resolution Displays," Grace Hopper Celebration of Women in Computing (Poster presentation)

**Other Reports**
Bradel, L., Endert, A., North, C. "Benefits of Co-located Collaboration on Large, High-Resolution Displays." Grace Hopper Conference 2011. Portland, Oregon. November, 2011. (Poster)

Bradel, L., Andrews, C., Endert, A., Vogt, K., Hutchings, D., North, C. "Space for Two to Think: Large, High-Resolution Displays for Co-located Collaborative Sensemaking." Technical Report TR-11-08. Computer Science, Virginia Tech

**Research Assistants**
Current Graduate Students
Lauren Bradel (Virginia Tech, Computer Science)
Undergraduate students
Katherine Vogt (Elon University, Computer Science)

**Follow-Up on Funding/Related Projects**
NSF REU (Research Experience for Undergraduates). Dr. Scott McCrickard (summer 2010, summer 2011)
NSF Equipment Grant for Large, High-Resolution Displays. Dr. Chris North (summer – fall 2010)

**Other Outcomes/Impacts**
We identified user roles that may develop in co-located collaborative analysis on large, high-resolution displays (information finder a.k.a. forager and synthesizer a.k.a. sensemaker). We found that design differences in analytical tools greatly impact the way users interact with the space as well as each other, resulting in different analytical strategies and levels of collaboration. Large displays proved to be beneficial to the collaborative analysis process by allowing the participants to spatially arrange data and partition the display into territories. Multiple input devices allowed the pairs of users to interact simultaneously with the display, but we found that they primarily used this functionality to conduct individual investigations instead of working on a joint task. Finally, we found that the spatial arrangements of documents produced more common ground between participants than those who did not arrange information spatially on the large display.

These findings have been used to improve design decisions regarding visual analytics systems used on large, high-resolution displays for collaboration. Our results will be published through

the INTERACT conference (Lisbon, Portugal, September 2011). We are continuing to analyze the data collected and anticipate submitting additional papers on this research.

## IV. Education Programs and Outcomes

During period two, VACCINE's educational initiatives spanned the entire career development pipeline ranging from K-12 programs through undergraduate and graduate level work, to professional education and training programs. The VACCINE mission is to educate current Homeland Security stakeholders and the next generation of talent and enable them to make effective decisions from the mass of multisource, multimedia data they will face in their careers.

VACCINE educational efforts targeted four main areas in Period two: minority serving institution (MSI) collaboration, K-12 projects, undergraduate and graduate programs, and professional training. In addition to the established programs, VACCINE focused on many outreach activities such as attendance at conferences and presentations to groups in order to promote VACCINE efforts and get support and participation for VACCINE year three and beyond. The education group has also submitted several proposals to further increase funding opportunities and future program development. Throughout all of these efforts, VACCINE and CCICADA have maintained a collaborative relationship with monthly conference calls, joint attendance and presentations at conferences and shared development of material and content.

During the 3rd quarter of period two, Marti Burns, VACCINE's Assistant Director of Education, pursued another career opportunity and left VACCINE. This left a void in the coordination of education-related activities. The VACCINE Directors have hired a new Assistant Director of Education in the first quarter of period 3. It is anticipated that our Education programs will be fully coordinated by the 3rd quarter of period 3.

### Minority and Underserved Programs

### Team

VACCINE:
*Purdue University:* David Ebert, Marti Burns
*University of Houston-Downtown:* Richard Alo, Erin Hodgess
*Florida International University:* Shu-Ching Chen
*Jackson State University:* Loretta Moore, Jacqueline Jackson
*Morgan State University:* Tim Akers
*Navajo Technical College:* Mark Trebian
CCICADA:
*Rutgers University:* Midge Cozzens

A significant amount of MSI educational opportunities are associated with specific research problems. For instance, in Section V, Research Problem Descriptions, UHD and FIU can be found to have provided expert guidance to their graduate students. It is within these projects that learning occurs, theses written, and PhD's formed. In lieu of being redundant, those projects will not be addressed here. From an overview perspective, these are the general accomplishments that VACCINE oversaw within its MSI education programs.

- Three Jackson State University students (Chicora Chandler, Jotham Greer, and Ashley Rhodes) participated in summer internships at Purdue and were integrated into the established Visual Analytic Summer Program based on the Summer Undergraduate Research Foundation framework.

- The VACCINE team won awards in 5 categories in the VAST Challenge 2010.

- University of Houston - Downtown Professor Erin Hodgess and graduate student Laura Sanchez worked with data from the Purdue University Student Health Center and looked at occurrences of various diseases. They developed an interactive package with the R statistics package to be used by medical personnel. Tools from statistical quality control and time series analysis are employed. Forecasts can be made in order to schedule personnel appropriately. They also collaborated with Purdue VACCINE and faculty on the 2010 VAST Challenge competition and this interschool team won a prize.

- In August, 2010, we announced a new position, Assistant Director of MSI Education, Richard Alo, from the University of Houston-Downtown. In this unique position, he has strategic oversight for development and implementation of VACCINE outreach initiatives involving minority serving institutions.

- 2010 ADMI/A4RC Conference: The 2010 ADMI/A4RC conference was held on April 8-11, 2010 in Jackson, MS. VACCINE was spotlighted in two conference sessions and the evening Social Networking Event. Undergraduate students Chicora Chandler and Jotham Greer participated in the conference session that spotlighted VACCINE's research opportunities for students.

- Maureen Biggers, from Indiana University, was involved in the SACNAS conference poster/brochures, affinity pod training and the monthly teleconference coaching sessions, STARS conference poster, ADMI/AARCS conference in April 2010, and the selection of student participants. She flew to Jackson State University and met face to face with each student in the pod to talk about how thing were evolving as they got ready to move to Purdue for a summer of research.

- Jackson State University conducted a special topics visual analytics course. The course was offered during the spring semester of 2010 for graduate special topics in Visual Analytics. Eight Computer Science students registered for the course. Dr. Ebert taught the first lecture and gave an introduction to Visual Analytics. The book used for the course was "Illuminating the Path: The Research and Development Agenda for Visual Analytics". Students reviewed papers and gave oral presentations on various topics from the book. For the term project, students downloaded/evaluated a visual analytics application, presented their findings and created a technical report of the work.

- VACCINE has added the Navajo Technical College as a funded partner beginning in Year 2. Dr. David Ebert and Dr. Richard Alo traveled to NTC to meet with faculty and executives of the school to develop a framework whereby VACCINE can:

  1. Promote 'best practices' to assist in NTC retention, recruitment and motivational activities for their new four year computing degree programs.

  2. Assist in further development of Cyber Infrastructure-enabled distributed research and education network providing visualization education and research opportunities to NTC faculty and students.

  3. With above, proceed to develop an Affinity Research Pod (ARP) to join the research of the ARPs within the VACCINE group at Center for Computational Science and Advanced Distributed Simulation (*CCSDS*/ University of Houston Downtown UHD.

- Near the end of period 2, VACCINE added Morgan State University as a partner school. Morgan State University is an accredited four-year HBCU that received a Department of Homeland Security (DHS) Scientific Leadership Bridge Award (SLBA) for Minority Serving Community Colleges. The SLBA Program supports DHS S&T's commitment to the development of a well qualified HS-STEM workforce that reflects the population of the United States. Planning began in period 2 to host 15 interdisciplinary students from Morgan State for a 2 week summer internship focused on visual analytics and homeland security. The internships will take place in June 2011.

## A. Jackson State Research Pod

**Team**

*Jackson State University:* Loretta Moore, Jacqueline Jackson
*Purdue University:* David Ebert, Ross Maciejewski

**Abstract**

The objectives of this project were to continue to: (1) raise awareness of VACCINE at Jackson State University; and (2) train undergraduates in research methodology through exposure to a real world design problem.

**Publications**

Chandler, C., and Greer, J. 2011. Emergency Pocket Placard: Providing Emergency Response Information via Mobile Application. In *Proceedings of the 2011 Association of Computer and Information Science/Engineering Departments at Minority Institutions Symposium on Computing*, Clemson University, Clemson, South Carolina, April 14-16, 2011.

**Presentations**

Chicora Chandler, Poster Presentation: Emergency Pocket Placard: Providing Emergency Response Information via Mobile Application, STARS Celebration Conference 2010, Champions-Gate, FL, August 10, 2010

Jotham Greer, Poster Presentation: Emergency Pocket Placard: Hazardous Materials Placard Identification Visual Analysis Tool for First Responders, STARS Celebration Conference 2010, Champions-Gate, FL, August 10, 2010.

Ashley Rhodes, Poster Presentation: Emergency Pocket Placard: Image Processing of HazMat Placards Using MatLab, STARS Celebration Conference 2010, Champions-Gate, FL, August 10, 2010.

Chicora Chandler, Poster Presentation: Emergency Pocket Placard: Providing Emergency Response Information via Mobile Application, 2010 AMIE conference, Jackson, MS, September 9, 2010.

Jotham Greer, Poster Presentation: Emergency Pocket Placard: Hazardous Materials Placard Identification Visual Analysis Tool for First Responders, 2010 AMIE conference, Jackson, MS, September 9, 2010.

Ashley Rhodes, Poster Presentation: Emergency Pocket Placard: Image Processing of HazMat Placards Using MatLab, 2010 AMIE Conference, Jackson, MS, September 9, 2010.

Jacqueline Jackson, Introduction to Visual Analytics for Command, Control, and Interoperability Environments: A Special Topics Course, Fall Meeting of the Mid-Southeast Chapter of the ACM, Gatlinburg, TN, November 11-12, 2010.

Chicora Chandler and Jotham Greer, Emergency Pocket Placard: Providing Emergency Response Information via Mobile Application, 2011 Association of Computer and Information Science/Engineering Departments at Minority Institutions Symposium on Computing, Clemson University, Clemson, South Carolina, April 14-16, 2011.

**Reports**

Jackson, J. 2011. 2010. Introduction to Visual Analytics for Command, Control, and Interoperability Environments: A Special Topics Course. In *Proceedings of the 2010 Fall Meeting of the Mid-Southeast Chapter of the ACM*, Gatlinburg, TN, November 11-12, 2010. (Published Abstract)

**Research Assistants**

Ashley Rhodes, Jackson State University, Department of Computer Science, (supported: April 2010 – August 2010)

Two undergraduate students also supported

**Other Outcomes/Impacts**

This MSI collaboration has provided students and faculty with numerous professional development experiences. As a result of this project, (1) three students (Chicora Chandler, Jotham Greer, Ashley Rhodes) completed summer internships at Purdue during the Summer of 2010 and showcased their work on the Emergency Pocket Placard at three conferences: STARS Celebration 2010, AMIE 2010, ADMI/A4RC 2011; (2) Dr Jacqueline Jackson spent a week at Purdue during Summer of 2010 meeting with students, faculty and the project director to facilitate JSU's collaboration with VACCINE; (3) a special topics course in Visual Analytics was developed and taught at Jackson State University at both the undergraduate and graduate level; and (4) VACCINE and Visual Analytics were highlighted in a newly developed course on Computational Thinking. Accomplishments #3 and #4 provided approximately 40 students' exposure to VACCINE and Visual Analytics.

Dr. Jacqueline Jackson presented JSU's Visual Analytics Course at the 2010 ACM Mid-Southeast Conference held in Gatlinburg, TN. The students, Chicora Chandler and Jotham Greer, received a 1st place prize for their oral presentation "Emergency Pocket Placard: Providing Emergency Response Information via Mobile Application" at the ADMI/A4RC 2011 Symposium. Their work was also accepted for publication in the conference proceedings.

## K-12 Program

*Due to fiscal constraints and to accommodate DHS wishes, the VACCINE K-12 program has been phased out at the end of period 2. The HS-STEM program will continue to have involvement with K-12.*

## Team

VACCINE:
*Purdue University:* David Ebert, Marti Burns
*Discovery Middle School:* James Howard, Kent Mikel
*Triton Central High School:* Steve Elder

CCICADA:
*Rutgers University:* Midge Cozzens

## Abstract

Part of the VACCINE mission in period 2 was to educate the next generation of talent through the VACCINE K-12 programs which include; the development of middle school and high school teacher workshops, a high school teacher tutorial for the high school visual analytic module, and the development of middle school modules. In addition, outreach for the K-12 programs included; the attendance at conferences and events, as well as classroom and organization presentations.

**Milestones**

| Milestones for K-12 Teacher Workshops | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | **Quarterly** | | | |
| Meet with Director of Discovery Park K-12 Education to discuss DP collaboration – Workshop coordination | X | | N/A | N/A |
| Quarterly Planning Meetings | X | | N/A | N/A |
| Discuss Collaboration with Rutgers – monthly conference calls | X | | N/A | N/A |
| Workshops offered (2) June 3 & 4 | X | | N/A | N/A |
| Workshop Evaluations Completed and Summarized | | X | N/A | N/A |

**Program Outcomes (Results, Accomplishments, Conclusions)**

Visual Analytics for Command, Control and interoperability Environments, VACCINE, hosted two one-day Middle School/High School Teacher Workshops on June 3 and June 4, 2010.  The workshops were held at the new Discovery Learning Research Center in Discovery Park on Purdue's West Lafayette campus. Twenty nine teachers from a variety of disciplines (Math, Science, Social Studies, and English) and schools from all over Indiana participated in the workshops.  The facilitators and participants were very pleased with the attendance and outcomes of the workshop and hope to initially replicate these efforts through CCICADA and University of Houston channels.

Workshop Focus:  Introduce visualization and data analysis to middle school and high school teachers through homeland security examples, curriculum development and hand-on learning modules.

Workshop Goals:

1. Define Visual Analytics (VA) and teach participants how they can incorporate VA into their curriculum to enhance learning.
2. Explain how to identify misinformation within graphics, good & bad visualizations and good and bad data.
3. Show participants how to find online resources for further explorations.
4. Determine outstanding teacher curriculum needs and ideas for future workshops.

To eliminate a lot of paper, we supplied the participants with a link to a webpage where the workshop materials and resources are posted.  Teachers can access the materials they need as well additional materials can be added and updated and as needed. **Workshop Materials can be accessed via:** http://pixel.ecn.purdue.edu:8080/~purpl/VisualAnalyticsWorkshop/index.html

VACCINE had support from Discovery Park, Wabash Valley Education Center and VACCINE graduate students.

Workshop Facilitators:

Marti Burns – Asst. Director of Engagement & Education, VACCINE
Steve Elder – High School Government / History, Triton Central High School
Chris Foster – Director PK-12 Education, Discovery Park

James Howard- 7th Grade Social Studies, Discovery Middle School
Nick Klosterman – Graduate Student, VACCINE
Kent Mikel - 7th Grade Science, Discovery Middle School

Workshop Assistance:

Tami Hicks – Wabash Valley Education Center
Valerie Lawless – Engagement Operations Manager, Discovery Park Administration
Stefanie Cox – Secretary, VACCINE
Lana Rice – Secretary, P-14 STEM Program, Discovery Park

Vaccine Tools Demonstrations:

Tim Collins – Managing Director, VACCINE
Nitin Khana – Rosetta Phone
Sohaib Ghani – COE Explorer
Abish Malek – VALET

Evaluations:

Each attendee filled out an evaluation form at the end of the day. Overall, the participants agreed that it was a valuable workshop full of a lot of information. Twenty of the participants indicated they would definitely like to work with VACCINE on future curriculum development. By the end of the day all participants were able to define visual analytics and begin to explore ways they could use visualization materials in their classrooms.

Evaluations from the participants also included the following comments:

Most helpful pieces of the workshop:

Very informative
Good to know VACCINE exists at Purdue
Lots of good resources and of good ideas
Understanding practical uses of technology
Good to hear ideas from other teachers
Liked the Rosetta phone demonstration the best

Challenges / suggestions for improvement:

Divide teachers by discipline for small group discussion
Too much information for one day – hard to process it all
Most participants indicated they would have liked to have more time not only to explore online resources, but to share ideas among other teachers in small groups.

What they would like to see more of in the future:

Hands on training
More info about how VA can be used in the classroom

Updates on additional tools and resources, more links and ideas
Notification of VACCINE tool updates and when tools are available for public use
Step by step instructions for using resources
Information about how teachers can use visualization for themselves with student analysis
Follow up workshop next year to see how teachers have used these materials

## Undergraduate and Graduate Programs

### A. Undergraduate Visual Analytics Summer Program (VASP)

**Team**

VACCINE:
*Purdue University:* David Ebert, Marti Burns, Ross Maciejewski
*Jackson State University:* Jacqueline Jackson, Loretta Moore

**Abstract**
As a part of Purdue's long standing Summer Undergraduate Research Fellowship Program, VASP is a hands-on summer learning program that introduces students to visual analytics research projects. All students present their research at the end of the summer. VASP 2010 hosted 3 Jackson State Students and 8 additional students from May 24, 2010 to August 6, 2010.

The summer program is an intensive 11-week program where students participate in research activities and attend professional development seminars for approximately 40 hours a week. The professor outlines the student's research schedule and appoints a graduate student mentor for the undergraduate.

The priority of the VACCINE Center with our VASP program is to recruit highly qualified students to graduate school, with particular emphasis on underrepresented populations. The VASP program provides students across engineering, science and technology disciplines with an intensive visual analytic research experience, allowing them to work closely with graduate students and professors in their respective schools while working on important homeland security problems.

The interdisciplinary aspect of the projects allows students to learn and work across other disciplines while still applying the concepts and skills from their own programs. This setting provides undergraduate students with an avenue to perform research in an academic environment while exploring future graduate study options.

VASP Program Benefits

Paid, hands-on research under the guidance of a faculty member and a graduate student
Weekly seminars on research methodology, graduate school, and professional development
Student poster presentations
Social activities with other SURF students
End of summer banquet

## Project Assignments

| Student | Project |
|---------|---------|
| Jotham Greer | Rosetta Phone - Gang Graffiti |
| Skye Chandler | Rosetta Phone - Droid Platform |
| Vishal Gala | COE Explorer |
| Yicheng Guo | COE Explorer |
| Tinghui Zhou | Mobile Evacuation and Bristle map rendering within spatial displays |
| Jing Li | VALET |
| Ahmad Razip | Kimberly Clark Point of Sale |
| Charles Smith | VAST Contest |
| Di Huang | VAST Contest |
| Shriphani Palakodety | IN-SPIRE analysis of COE database |
| Daudi Muhamed | H-PAC into PlumeVis |

## Milestones

| Milestones for VACCINE Period 2 – Undergraduate Visual Analytics Summer Program | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| 2010 Program applicant selection continues | X | | | |
| VASP summer internship | | X | X | |
| Evaluate 2010 program, suggestions/corrections for 2011 | | | X | |
| Outreach to partnering schools for applications | | | | X |
| 2011 Program applicant selection begins | | | | X |

## Program Outcomes (Results, Accomplishments, Conclusions)

Other outcomes/impacts

Selections have been made for the VASP 2011 program and 4 students will participate with VACCINE summer research projects. VACCINE and CCICADA have collaborated with their partner schools for the advertising of the program.

**B. HS-STEM Career Development Program**

**Team**

VACCINE:
*Purdue University:* David Ebert, Marti Burns

**Abstract**
The HS-STEM Career Development program is a competitive program funded through DHS that was established at Purdue in 2007 under David Ebert. Purdue offered fellowships to seven graduate students in 2007, three graduates and one undergraduate student in 2008, and three graduate students in 2010. VACCINE has applied for a continuation of funding through a proposal for the HS-STEM 2011 program.

**Purdue University**
**HS-STEM Career Development Program Activities**
**Annual Summary for 2010**

HS-STEM Graduates:

**David Blunck** graduated in May, 2010 with his PhD in Mechanical Engineering and went to work for Air Force Research Laboratories to meet his service requirement.

**Karla Combs** completed her PhD in Biological Sciences in May, 2010. Karla completed her year of service at the Illinois State Department of Health in the Emergency Preparedness section. She worked on H1N1 planning for the State of Illinois and assisted in planning and executing full-scale exercises to test these plans.

**Laura Hughes** finished her PhD in Biological Sciences in August, 2010. Laura is now working at the Emory Vaccine Center to complete her year of service.

**Nwokedi Idika** completed his PhD in Computer Science in August, 2010. He is currently working at MIT Lincoln Labs to complete his year of service. Nwokedi has the distinction of being the first African American PhD in Computer Science from Purdue University.

**Bryan Sims** graduated in December, 2010 with a PhD in Nuclear Engineering and is currently employed by Idaho National Labs for his year of service.

Monthly HS-STEM Student Lunches:

Networking for students and advisors.
One student presents their research and activities each month.

Information sharing: poster/presentation opportunities, campus opportunities, DHS opportunities, etc.

Presentations / Posters:
Students are encouraged to do presentations and present posters during campus activities as well as off site conferences (see individual summaries for specifics).

DHS UP Conference in October 2010
Required for all recipients to attend. Since we did not admit a new class of students in 2009, we did not have any students attend in 2010 but will have several students attending in 2011.

Reporting:
Student and advisor summary reports due twice each year, May and January. Annual Report to DHS University Programs – submitted January 2011

Goals for 2011:
Submitted proposal for HS-STEM Career Development Program 2011 and hope to bring a new group of graduate students on board in January 2012.

Plan programs/activities to help integrate new students with current students; use current HS-STEM students to assist in recruiting new students where appropriate.

Encourage collaboration among HS-STEM I, II and III students and advisors.

Increase knowledge of internships opportunities for undergraduates and job placement for both undergraduates and graduates.

Expand relationships with other Centers of Excellence and their HS-STEM programs.

Continue HS-STEM Career Development Program integration with K-12 efforts by having HS-STEM students present in local middle/ high schools to present their research. Post these presentations on the VACCINE website for access by other teachers.

HS-STEM students hope to sponsor a speaker of campus wide interest during 2011.

**Milestones**

| Milestones for VACCINE Period 2 – HS-STEM Career Development Program | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| Present in High School Classrooms | X | X | X | X |
| Attend monthly UP conference Calls | X | X | X | X |
| Monthly lunches ( highlight one student/month) | X | X | X | X |
| Job recommendations / placements for students | X | X | X | X |

**Program Outcomes (Results, Accomplishments, Conclusions)**

Other Outcomes/Impacts
Fourteen students at Purdue University have been funded through the HS-STEM Career Development program. The program is led by Dr. David Ebert, Director of VACCINE and

managed by Marti Burns, Asst. Director of Engagement and Education. To date, seven students have graduated with PhDs and have gone on to continue work in homeland security fields.

## C. Graduate Course Program

VACCINE:
*Georgia Institute of Technology:* John Stasko

### Abstract

This component of the education mission involves graduate programs and education at the VACCINE Center. One specific goal of this project was to promote and coordinate graduate courses and programs at the VACCINE member universities and beyond. The Center gathered and maintained a list of related courses already being taught at our universities. Also, together with colleagues from the FODAVA Center and NVAC, in 2009/2010, the team created and publicized a "Body of Knowledge" for visual analytics.

### Milestones

| Milestones for VACCINE Period 2 – Graduate Programs | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| Finalize directory of visual analytics courses and education programs | X | | | |
| Plan an education workshop about visual analytics courses | X | | | |
| Workshop | | X | | |
| Begin work on developing a Masters degree Area of Specialization | | | X | X |

### Program Outcomes (Results, Accomplishments, Conclusions)

Other Outcomes/Impacts

The team gathered an initial list of courses and education programs at our member universities, but they have not made this into web pages for the VACCINE website yet. In the current educational climate of limited budgets and expansion, the team decided not to spend time developing a new visual analytics course or Masters degree. Instead, the team proposed that a Masters Area of Specialization be developed as a pilot project at both Purdue and Georgia Tech.

On Monday August 30, in conjunction with the Visual Analytics Community Consortium meeting (http://nvac.pnl.gov/meeting_2010/), we held a one-day workshop on visual analytics education at the University of Maryland in College Park, MD. This workshop brought together those currently teaching courses closely related to visual analytics and those who were interested in doing so.

The goals of this workshop was to share lessons learned and best practices in teaching visual analytics, and to develop an example curriculum that will be vetted with the broader academic community.

The workshop featured interactive presentations by several people who have developed or taught visual analytics related courses. It included breakouts and group discussions on particular topics related to visual analytics education.

**D. VAST Challenge 2010**

**Abstract**

The VAST Challenge is a participation category of the IEEE VAST 2010 Symposium (part of VisWeek 2010). The VAST 2010 Challenge continued in the footsteps of the VAST 2009 Challenge, VAST 2008 Challenge and the 2007 and 2006 contests with the purpose of pushing the forefront of visual analytics tools using benchmark data sets and establishing a forum to advance visual analytics evaluation methods. Another goal of the VAST Challenge is to speed the transfer of VA technology from research labs to commercial products, and increase the availability of evaluation techniques.

In order to provide more opportunities for increased participation, the VAST Challenge is comprised of an overall Grand Challenge as well as three smaller Mini Challenges.

Mini Challenge 1: Text Records - Investigations into Arms Dealing
Mini Challenge 2: Hospitalization Records - Characterization of Pandemic Spread
Mini Challenge 3: Genetic Sequences – Tracing the Mutations of a Disease Grand Challenge: Combines all data sets. Investigating any possible linkage between the illegal arms dealing and the pandemic outbreak.
Teams may enter one or more mini Challenges independently of entering the Grand Challenge. ALL teams submitting an entry to a VAST Challenge will be invited to discuss their work during a challenge workshop.

Entries are judged on both the correctness of the analysis (based on the availability of ground truth) and the utility of the tools in conducting the analysis. Participants have several months to prepare their submissions.

| Milestones for VACCINE Period 2 – VAST Challenge | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | **Quarterly** | | | |
| Full datasets available (end of March) | X | | | |
| Submission deadline, June 29, 2010 | X | | | |
| Results returned to participants | | X | | |
| Deadline for award winners to submit 2-page summary for inclusion in the proceedings | | X | | |
| VAST Conference – Awards, October 2010 | | | X | |

**Program Outcomes (Results, Accomplishments, Conclusions)**

Mini Challenge 1: Text Records - Investigations into Arms Dealing
      Award – Good Support for Data Ingest
      Georgia-Institute of Technology

      Award – Outstanding Interaction Model
      Simon Fraser University

Mini Challenge 2: Hospitalization Records – Characterization of Pandemic Spread
      Award – Support for Future Detection

Purdue University

Mini Challenge 3: Genetic Sequences – Tracing the Mutations of a Disease
Award – Excellent <u>Process Explanation</u>
Georgia Institute of Technology

Grand Challenge: Combines all data sets.
Award – Excellent Student Team Analysis
Simon Fraser University

**Professional Training**

**A. RECONNECT Workshop**

**Team**
VACCINE:
*Purdue University:* Tim Collins, Marti Burns
*Georgia Institute of Technology:* John Stasko
CCICADA:
*Rutgers University:* Midge Cozzens

**Abstract:** This ongoing CCICADA/VACCINE Summer Reconnect Workshop exposes faculty teaching undergraduates to the role of the mathematical sciences in homeland security and provides an opportunity to researchers in government or industry to learn about recent techniques in data analytics. Topics are presented in a weeklong series of lectures and activities; participants are involved in both research activities and in writing materials useful in the classroom or to share with their colleagues. Participants may develop materials for publication in either the CCICADA Technical Reports or the Educational Modules Series published by the DIMACS Center at Rutgers University. The 2010 RECONNECT workshop was held at the University of Southern California on June 6-12, 2010.

Automatic identification and extraction of desired information from natural language text is increasingly used as a way to improve general purpose search and has a range of applications in medical informatics, business applications, and for the intelligence community. The input is one or more texts in the domain in question, and the output is a database containing just the desired fields of information, extracted from the source material and formatted appropriately. Information extraction (IE) techniques have been developed since the early 1980s, and include finite state technology, pattern-based extraction, and appropriate machine learning methods. This week-long workshop will take participants from the early, simpler, methods through the modern ones, and will include theoretical and practical topics as well as hands-on exercises using software packages. The material is very relevant to the undergraduate classroom and to many applications. The lecturers are renowned experts in the various aspects of IE and its visualization, and have a long history of giving informative, engaging, and fun lectures.

**Deliverables (Milestones)**

| Milestones for VACCINE Period 1 – Professional Training | 04/01/10 – 03/31/11 | | | |
|---|---|---|---|---|
| | Quarterly | | | |
| Discuss Professional Training opportunities with Rutgers / University of Houston DT /partner schools | X | X | X | |
| John Stasko teacher in Rutgers RECONNECT summer program | X | | | |
| Attend IN District 4 first responders meetings for opportunities/ needs | | | X | |
| Design program for VACCINE period 3 | | | | X |

**Program Outcomes (Results, Accomplishments, Conclusions)**

CCICADA/VACCINE Summer Reconnect Workshops exposed faculty teaching undergraduates to the role of the mathematical sciences in homeland security and provided an opportunity to researchers in government or industry to learn about recent techniques in data analytics. RECONNECT 2010 was held at USC. Midge Cozzens and John Stasko taught at RECONNECT during the summer 2010.

**B. JIBC Workshop**

Justice Institute of British Columbia and VACCINE collaborative workshop enabling visual analytics and precision information environments for first responders.

**Team:**
*Simon Fraser University:* Brian Fisher
*Justice Institute of British Columbia:* Murray Day, Bob Walker, Jack McGee
*Purdue University:* Tim Collins, David Ebert

**Abstract:** VACCINE is a true international collaboration. Our partners in Canada are critical to our success. VACCINE's strategy is predicated on building strong educational programs. We intend to educate the next generation of homeland security professionals to harness the power of visual analytics and advanced computational tools to enable them to make effective decisions from the mass of multisource, multimedia data they will face in their careers. We will develop innovative educational techniques for visual analytics and decision making. We will continue to create the leading international repository and resource for visual analytics educational and training materials. Through our partnership with Justice Institute of British Columbia (JIBC), we will develop and provide professional training and educational development materials and opportunities for homeland security professionals ranging from local emergency responders to state and federal officials.

The JIBC is Canada's leading public safety educator - a dynamic, post-secondary institution recognized nationally and internationally for continuous improvement and innovative education in the areas of justice and public safety. The Institute was established in 1978, and is currently comprised of eleven Academic departments and a number of supporting administrative departments. In addition, six regional campuses allow students the flexibility to study closer to home.

As part of its commitment to innovation and continuous improvement, the JIBC is also involved in applied research that spans the continuum of safety, from prevention to response and recovery. The goals of the JIBC's research initiatives are to inform public policy, enhance curriculum, contribute to public education and maximize the benefits of learning technologies and methodologies.

**Technical Approach**

On September 20 and 21, 2010, VACCINE, with its partner JIBC, held a workshop in support of visual analytics education and the first responder. It is imperative that public safety personnel acquire the requisite knowledge in visual analytics that will enable them to understand events, generate plans, execute strategies, and make tactical and operational decisions. Therefore, the overarching purpose of the workshop was to define the requirements that will lead to developing academic programs and relevant technology utilizing visual analytics in the public safety environment. This was accomplished through a series of lectures, panel discussions, and breakout sessions. All sessions of the workshop were recorded and made available on DVD.

Goals of the Workshop

Identification of public safety academic requirements related to visual analytics and precision information environments.

Identification of collaborative research areas among VACCINE/JIBC faculty.

Demonstration of the JIBC Simulation Center ExPod capabilities and identification of a framework to build a cross-border exercise Center of Excellence.

Showcasing of VACCINE technology to public safety officials.

Deliverables (Milestones)

| Milestones for VACCINE Period 2 – Professional Training | 04/01/10 – 03/31/11 Quarterly | | | |
|---|---|---|---|---|
| Weekly planning meetings | X | X | | |
| Travel to JIBC. Planning and discussion of agenda. | | X | | |
| Agenda development | | X | | |
| Workshop | | X | | |
| Ongoing discussion to incorporate VACCINE technology into JIBC Masters program (In development) | | | X | X |

**Program Outcomes (Results, Accomplishments, Conclusions)**

Program materials can be found at http://www.purdue.edu/discoverypark/vaccine/jibc.php.

Work is now underway to incorporate VACCINE technologies into the JIBC Masters program being developed. Pilot project underway to integrate ExPod with CrimeViz and include this into the Masters curriculum.

## C. ULearning Platform

**Team**
*Purdue University:* David Ebert, Tim Collins
*Captico, LLC:* Regina Perkins, Corrie Davidson

**Abstract:** The Department of Homeland Security Science and Technology Directorate (DHS/S&T/CCI) focuses on the development of education and training tools, as well as collaboration space, for the emergency response community. Captico has partnered with VACCINE (a DHS Center of Excellence) at Purdue University to develop a web-based uLearning prototype for DHS/S&T/CCI to serve the emergency response community. The prototype leverages content such as videos, presentations, narratives and links to share expert and practitioner-defined content across the emergency response community. The objective of this effort is to develop an initial short, uLearning (user learning) tool that incorporates video, text, images, or audio of new technologies and how new technologies can be used in operational scenarios.

**Deliverables**

Prototype uLearning Tool
Assessment of uLearning tool with small set of managers/users
Plans/approach for broader uLearning applications (e.g., other technologies and end-users)

**Project Outcomes (Results, Accomplishments, Conclusions)**
Captico and Purdue have identified and prioritize learning needs within the emergency response community that may be best supported via the uLearning tool. By gathering input from first responders utilizing various methods, the uLearning tool prototype was developed to incorporate specific features and capabilities found most useful by the emergency response community.

Based on availability of funding, the plan is to target the state of Indiana as a large pilot project and give those first responders supported access to the uLearning Platform in exchange for feedback to further the development of the Platform.

In addition to formal training materials, first responders regularly create in-house videos, documents, and presentations to educate themselves and their fellow team members. Some counties use private file sharing services to trade materials across departments. Others mail CDs with videos on them back and forth due to privacy concerns. Others upload content to YouTube and hope for the best. There is an astounding amount of educational content that is created and then wasted by limits on interoperability and a lack of an easy-to-use, approved, central repository. When presented with the uLearning Platform (particularly in person) first responders got excited about the opportunities to better share materials not just within their own Departments, but also with others. Volunteers and small departments especially are desperate for resources and eager to take advantage of the wealth of information that the uLearning Platform could grant them.

The complete ULearning pilot report can be found in Addendum 3.

# V. Partnerships and Outreach

## 1) List of Partnerships and Major Outcomes

VACCINE has been very active in both its domestic and international outreach efforts. We have built upon our existing relationships and established many new collaborative partners.

### 1) <u>VACCINE Public Safety Partnerships</u>

| | |
|---|---|
| **1. Indiana Public Safety Consortium**<br>   a. Tippecanoe County Sheriff's Office<br>   b. Tippecanoe County Emergency Management Agency<br>   c. Lafayette Police Department<br>   d. West Lafayette Police Department<br>   e. Purdue University Police Department | **2. Florida Division of Emergency Management's Region Seven**<br>   a. Florida International University<br>   b. Miami-Dade County<br>   c. Broward County<br>   d. Palm Beach County<br>   e. Monroe County |
| **3. UNCC Consortium**<br>   a. Charlotte Mecklenburg Police Department | **4. Washington State Consortium**<br>   a. Port of Puget Sound<br>   b. U.S. Coast Guard, Seattle Sector |
| **5. Penn State Law Enforcement Consortium**<br>   a. Harrisburg Police Department | |

In 2010, VACCINE created the Indiana Public Safety Consortium which is comprised of local Indiana law enforcement and emergency management agencies. This Consortium was established with the goal of connecting researchers with first responders. Memorandums of Understanding have been established with all agencies which will permit the sharing of data, information and materials needed to conduct projects with the goals to coordinate and to improve public safety's analysis of data to promote improved operations in the law enforcement, fire, and emergency management and EMS fields. The group meets monthly to discuss tools that are available, as well as to brainstorm tools that are needed. VACCINE envisions the Consortium as a resource that can be utilized by not only VACCINE but also DHS S&T.

<u>Technology Deployment</u>

VACCINE's mission is to create tools and technologies for the first responder community. Understanding that most first responders do not sit behind a desk during their normal working day, we have created a suite of tools to literally put into the hands of our stakeholder. Our suite of tools includes:

**VALET (Visual Analytics Law Enforcement Toolkit)**

The Visual Law Enforcement Toolkit (VALET) provides law enforcement agencies with a suite of analytical tools coupled with an interactive visual interface for data exploration and analysis, forecasting, planning, and management. This system includes linked views and interactive displays that spatiotemporally model criminal, traffic and civil (CTC) incidents and allows officials to observe patterns and quickly identify regions with higher probabilities of activity.

Our toolkit provides analysts with the ability to visualize different types of data sets (census data, daily weather reports, zoning tracts, prominent calendar dates, etc.) that provide an insight into correlations among CTC incidents and spatial demographics. In the spatial domain, we have implemented a kernel density estimation mapping technique that creates a color map of spatially distributed CTC events that allows analysts to quickly find and identify areas with unusually large activity levels. In the temporal domain, reports can be aggregated by day, week, month or year, allowing the analysts to visualize the CTC activities spatially over a period of time. Furthermore, we have incorporated temporal prediction algorithms to forecast future CTC incident levels within a 95% confidence interval. Such predictions aid law enforcement officials in understanding how hotspots may grow in the future in order to judiciously allocate resources and take preventive measures. Our system has been developed using actual law enforcement data and is currently being evaluated and refined by VACCINE's Public Safety Consortium of law enforcement agencies.

The VALET software is available on a PC, iPhone, or iPad and has been deployed to members of the VACCINE Public Safety Consortium including the Tippecanoe Co. Sheriff's Department, Lafayette Police Department, West Lafayette Police Department, and Purdue University Police Department. VALET provides advanced analysis capabilities, allowing analysts to develop and test hypothesis about criminal activities within various areas of their communities.

**MERGE (Mobile Emergency Response Guide)**

Hazardous Materials can react differently to stimuli and can cause problems in accidents and emergency situations. This fact makes them particularly dangerous to civilians and first responders. It is for this reason that the Emergency Response Guidebook was developed. This paper guide assists those in an emergency with the knowledge of how to handle Hazardous Materials. As one might expect, the guidebook is large and requires precious time to search an index to determine the best way to handle a particular hazardous material.

The MERGE system, shown below, is an electronic version of the guide with many new features and capabilities. These new capabilities include the use of image analysis methods to automatically determine the type of Hazardous Materials present based on an image taken of the sign/placard. MERGE has an easy to understand user interface to instruct an emergency responder or civilian in the proper way to handle a Hazardous Material emergency.

MERGE was developed at the request of our first responders. The software is underdevelopment for use on iPhone and Android mobile devices. We anticipate that this should be deployed by summer 2011. We have requests for use from the Lafayette, West Lafayette, Indianapolis, and St. Clair, MI Fire Departments.

**GARI (Gang Graffiti Automatic Recognition and Interpretation)**

Gangs are a serious threat to public safety throughout the United States. Gang members are continuously migrating from urban cities to suburban areas. They are responsible for an increasing percentage of crime and violence in many communities. According to the National Gang Threat Assessment, approximately 1 million gang members belonging to more than 20,000 gangs were criminally active within all 50 states and the District of Columbia as of September 2008. Criminal gangs commit as much as 80 percent of the crime in many communities according to law enforcement officials throughout the nation.

Street gang graffiti is their most common way to communicate messages, including challenges, warnings, or intimidation to rival gangs. It is, however, an excellent way to track gang affiliation and growth, or even sometimes to obtain membership information. The goal of this project is to use the knowledge gained from our work in mobile devices and applications and leverage it towards the development of a mobile-based system capable of image analysis. This system will provide an accurate and useful output to a user base through a database of gang graffiti images.

The image analysis includes obtainment of metadata (e.g., geoposition, date, and time) and the extraction of relevant features (e.g., color, shape) from the gang graffiti image. The information is sent to a server and compared against the graffiti image database. The matched results are sent back to the device where the user can then review the results and provide extra inputs to refine information. Once the graffiti is completely decoded and interpreted, it is labeled and added to the database.

GARI is under development for use on iPhone and Android mobile devices. We anticipate that this should be deployed by fall 2011. We have requests for use from the Lafayette, West Lafayette, Indianapolis, and Purdue Police Departments.

## 2) VACCINE DHS Operating Components

### a. United States Coast Guard

1) CgSARVA – Coast Guard Search and Rescue Visual Analytics

VACCINE has spent the past year building strong ties within the USCG. We have collaborated with them to build tools that can be adapted to real time situational areas. Most significant of these projects was a detailed model allowing a thorough assessment of all CG search and rescue cases in the Great Lakes. In the Coast Guard Search and Rescue Visual Analytics (CgSARVA) project, VACCINE presents a risk analysis system that enables the interactive visualization, analysis, and assessment of SAR missions completed by each USCG station.

The CgSARVA technology is deployed, and in use, at the USCG Atlantic Area Planning and Analysis Division. It will be deployed at the unit level during the summer of 2011.

The system enables the interactive analysis of trends, patterns, anomalies, and distribution of cases and associated sorties. Additionally, this assessment tool enables the determination of potential increase or decrease in risk with a reallocation of a resource; as well as known increases or decreases in the response time. The CgSARVA system features include the following:

- Linked calendar and line graph views for analyzing data patterns and distributions
- Interactive filter controls for assessing case distribution loads and locations
- Density estimation for hotspot generation
- Linked time slider controls for interactive temporal animation and exploration
- Interactive statistical summary tools for report generation

The CgSARVA technology is a vital component for analyzing risk assessment as Coast Guard staffing levels fluctuate due to budget changes, retiring volunteers, and a variety of other factors. This system provides managers and analysts with a suite of tools for analyzing the distribution of previous search and rescue cases and a methodology for understanding the risk, efficiencies, and benefits involved with reallocation or reduction of resources.

2) Analytical Visualization of the Boston PWCS Patrol Resource Allocation Environment (with CREATE) aka PROTECT-Port Resilience Operational / Tactical Enforcement to Combat Terrorism

This is a $160,000 contract under our Basic Ordering Agreement. Using the previous two years of patrol routes, we are developing statistical models for loiter point visits. Predictions of the loiter point visits will be computed and sent to the Coastguard to compare against actual visits during a time period in which the visits were unknown to the analysts. The predictive efficacy of this model will then be addressed, and determination on the ability of hostiles to predict patrol routes based on current patterns will be assessed. Sparse data will be handled by predicting the number of days between visits. Similarly in utilizing the PROTECT model, two years worth of patrol routes will be generated from the PROTECT model. Similar statistical models will be created and a predictive model will be adapted to the PROTECT routes. Predictions of future loiter point regions will be generated from PROTECT and sent to us and our model will now assess the ability of hostiles to predict patrol routes based on potential PROTECT routes.

An interactive visual analytics environment will be created that will overlay patrol routes and potential Risk Index Numbers (RIN) numbers for targets. This environment will be developed for both desktop and hand-held platforms with a first generation tool looking at historical patrol route logs and critical infrastructure overlays. The first generation tool will be deployed to a provisional set of members for analysis and feedback.

3) Analytical Visualization of the Port Arthur, TX Economic Impact Study (with CREATE)

This is a contract under our Basic Ordering Agreement. In conjunction with the resources allocation tool, we will also look at creating economic impact models that will show the effect that potential catastrophes (natural or manmade) will have. Our system will incorporate new models, showing the potential impact over the Port Arthur area utilizing census tract information, historical data, the output of the CREATE economic model and predictive values. Initial versions of the system will focus on specific threat scenarios as determined by the partnering institutions. Our analytic environment will allow views and analysis of economic impact over time, over space, and over geography to show linkages and causalities and enable a better understanding of the relationship of certain primary and secondary economic effects.

4) Analysis of Swimmer Drownings in the Great Lakes (Summer 2011 project)

USCG District 9 expressed concern that the number of lives lost has increased dramatically over the last summer. The number of SAR cases in D9 matches

closely to the number of cases in FY09 and is down from the 5-year average; however, the number of lives lost has increased. LANT-7 reviewed the SAR cases where lives were lost after CG notification to try and understand the root cause of the increase and determine if a helicopter response capability was a factor. Of the 23 SAR cases reviewed, only one had a request (from an supporting government agency) for a CG helicopter that could not be filled. The Atlantic Area Aircraft Scheduler indicated that D9 has maintained both Air Facilities (Waukegan and Muskegan) and their associated B-0 aircraft along with B-0 aircraft at both permanent air stations (Traverse City/Detroit). Many of the cases with lives lost were persons in the water, close to shore, who were found beneath the surface. LANT-7 recommends that further analysis (i.e. reviewing lives lost after CG notification to determine root cause and possible prevention measures) be considered before a plan of action to combat the increasing loss of live within D9 can be established.

b. **United States Secret Service**
Disposable Cell Phone Analysis

This research project is in support of the Secret Service. Disposable cell phones have become popular in the marketplace. These devices are attractive to terrorist and criminal perpetrators because they allow a purchaser to have access to pre-paid cellular time without having to submit extensive personal information. The devices are also attractive because they are inexpensive and easy to use.

Disposable cell phones are intended for use over a limited period of time. Therefore, these devices frequently have their connectivity functionality disabled, not allowing data syncing with a personal computer. When the devices are used in the commission of criminal or terrorist activity, LE investigators are faced with a challenging forensic examination problem.

Current forensic examination methods on these devices are laborious and time consuming, the phones must be disassembled and analyzed or investigators must take photos of the screen to present evidence that is resident on the device. Neither of these methods allow an examiner to view all of the data, easily parse the data or examine files on the questioned device.

This effort will focus on the demonstration and development of methods and tools that will allow an investigator to acquire call logs, contacts, pictures, videos and text messages from disposable cell phones.

### c. **Federal Emergency Management Agency**

1) Business Continuity Information Network

Studies have shown that businesses risk failure if they are unable to reopen quickly after a disaster. To reduce these risks and improve communications within the business community, VACCINE partner, Florida International University has developed the Business Continuity Information Network (pronounced "bee-kin"), a web-based service where local businesses, emergency management, and organizations that assist businesses can gather to share critical information and support continuity efforts before, during and after a disaster. The BCIN South Florida Release supports business recovery programs initiated by county participants including Broward, Miami-Dade, Palm Beach and Monroe County through their respective Emergency Management offices.

Available year-round as a public service, this business-to-business community network provides participating companies a tool to track their key employees and supply chain status, and locate needed recovery goods and services. The system facilitates professional organizations like Chambers of Commerce to assist their members, and helps government agencies to assess damage and prioritize recovery needs. BCIN reports, maps and shares critical up to-the-minute information about infrastructure conditions and recovery efforts by working with County Emergency Management Offices and major private infrastructure providers. BCIN will localize and tailor relevant information to business managers who will use this information to better assess an event's impact to their facilities, employees, suppliers and customers.

BCIN's B2B network, which includes companies from different industry sectors, business and trade associations and NGOs, helps businesses locate disaster recovery resources, products, and services while allowing them to collaborate on recovery efforts. Businesses can use BCIN to report their operational
status and assistance needs to county government and NGOs so these organizations can prioritize their relief efforts. Recovery dollars spent locally helps to keep local businesses open and mitigates employment loss. The BCIN South Florida Release supports business recovery programs initiated by county participants including Broward, Miami-Dade, Palm Beach and Monroe County through their respective Emergency Management offices. Private sector participants include Office Depot, Wal-Mart, the Greater Miami Chamber of Commerce and a variety of local businesses across industry sectors. BCIN is supported by the Department of Homeland Security, IBM and the National Science Foundation.

2) A Data Integration Framework for Enhancing Emergency Response Situation Reports with Multi-Agency, Multi-Partner Multimedia Data

Responders in the field (both public and private participants) are capturing footage, pictures and video, of the disaster area with mobile phones. This data is being collected but is not integrated in the incidence command systems where situation reports, incidence action plans, etc are being held. When these materials are entered into such systems, it's done so in a manual way which is time consuming and requires careful human review and management.

To support emergency management operations, Florida International University is using advanced data integration and visual analytics techniques to analyze situation reports, incident action plans, and other type of responder reports, pictures and video clips captured in the field and automatically augment/link the reports directly to relevant multimedia content obtain in the field. Data integration techniques will be used to identify disaster specific keywords and will be used to identify the documents.

d. **Customs and Border Protection**
Query Performance of the Cargo Database (2010)

Massive amounts of data about ocean borne cargo shipments are now gathered from disparate sources and organized into a unified analytic database. To be useful for daily operations, such a database needs to be able to rapidly respond to analysts' *ad hoc* queries. This scenario, interactive analytic queries in large datasets, is one that VACCINE has studied at Stanford, but we had not had the opportunity to analyze with data of the size and scope found in the cargo domain. Working with Barry Siegel and Greg Meckstroth at SSC-PAC San Diego, VACCINE conducted a study of an extract of unclassified data from the 5TB Cargo database. The Cargo database was interesting not only because of its size, but also because it collects data from over two dozen different sources.
The principal finding of the study is that while the current structure of the Cargo database is remarkably flexible in terms of its ability to absorb data from a wide array of sources, this generality comes with a significant cost in terms of query performance. For the set of queries VACCINE was given as part of the study, Stanford worked out optimizations which improve query performance by factors of 2x-100x over the original Cargo database. This result is not surprising as database organization and query performance are interrelated. Good database design can only be done with intimate knowledge of the queries. Although security and privacy reasons precluded this study from having access to any actual queries, our findings still provide useful guidance for those analyzing real queries to restructure the Cargo database.

Endorsement:
*Let me add my endorsement to the report "Query Performance of the Cargo Database" written by Justin Talbot and John Gerth. We are developing a massive cargo database for use by Project SCREEN, a DHS S&T project on cargo targeting. The findings in the report are extremely applicable and directly supports the forthcoming Project SCREEN Broad Area Announcement.*

*Barry Siegel*
*SPAWAR System Center - Pacific*

3) **Additional VACCINE State and Federal Partnerships**

VACCINE works closely with state and federal agencies to receive data relevant to our projects. VACCINE has a well-established process for entering into Memorandums of Understanding and None Disclosure Agreements to ensure protection of the parties and data involved.

- FEMA (Federal Emergency Management Agency)
- IICD (Infrastructure Information Collection Division)
- GMO (Geospatial Management Office)
- DHS internal Geospatial Working Group
- DHS University Programs
- DHS US-CERT Operations
- DHS Research Transition Working Group
- U.S. Department of State, Office of the Geographer
- National Institute of Justice
- National Institutes of Health
- National Science Foundation
- Department of Health and Human Services
- Pacific Northwest National Laboratory
- Argonne National Laboratory
- Oak Ridge National Laboratories
- Defence Research & Development Canada
- Indiana State Department of Health
- Illinois Terrorism Task Force

4) **VACCINE Corporate Involvement**
In 2010, the VACCINE Center and the Purdue Research Foundation established a unique VACCINE Corporate Affiliate Program to promote VACCINE's goals and initiatives. In broad outline, this program provides the framework to enable corporations to make tax deductible gifts to support research, education, and training activities for all VACCINE member institutions. The basic structure is through the establishment of a non-profit Limited Liability Corporation (VACCINE LLC) which

is managed by the Purdue Research Foundation and overseen by VACCINE's Leadership Board. The VACCINE Leadership Board advises the LLC on how to distribute donated funds to support VACCINE's goals. The corporate donations are true gifts, with no strings attached. However, there are some practical benefits to donors, including the following:

- Membership in the VACCINE, LLC is by invitation only. Member companies need to be complementary as opposed to competitive in nature.
- Corporate affiliates should be in a position to enter into separate research agreements directly with member institutions to develop proprietary technologies, contractually obligated deliverables, etc.
- Corporate affiliates will be in a position to gain early information about ongoing basic research projects at member institutions before publication.
- Corporate affiliates will have opportunities to identify and recruit students at member institutions.
- The Corporate Affiliates program will provide a forum for donors to explore research collaborations and sponsorships.
- Corporate donors will have opportunities for employed scientists to take sabbaticals to participate in relevant research projects at member institutions.

The founding member of VACCINE, LLC was Motorola Solutions. In addition to Motorola, the VACCINE team has been engaged with a number of companies including the following:

- Harris Corporation
- Kimberly Clark
- The Boeing Company
- CRGT
- Oculus Info, Inc.
- ArgonST
- Bank of America
- Siemens
- General Dynamics
- Hallmark
- BanField, the Pet Hospital
- Aegis
- Kx Systems
- NVIDIA
- Duke Energy

**VACCINE**

Visual Analytics for Command, Control, and Interoperability Environments
A U.S. Department of Homeland Security Center of Excellence

# CORPORATE VISIONARY PARTNER PROGRAM

The VACCINE Corporate Affiliate Program offers companies a unique opportunity to collaborate with VACCINE faculty, students, and researchers working at the forefront of visualization and analytics research. Members of the Program are given unparalleled access to new technologies as they move from the laboratory to the marketplace.

At VACCINE, we view research as a practical matter — placing additional emphasis on designing and building technologies and solutions that have real-world impact — saving lives. In many cases, this requires establishing partnerships with industry to translate new ideas and discoveries to products and services already in the marketplace.

## FOUNDATIONAL RESEARCH AREAS

**Interactive Visual Analytic Foundations**

- Context-aware Mobile Visual Analytics for Emergency Response
- Mobile Imaging, Rosetta Phone and Light-weight Visual Analytics for In-field Analytics

**Visually-Adapted Analytical Techniques**

- Cybersecurity Visual Analytics
- Visual Analytics for Investigative Analysis on Text Documents
- Multimedia Visual Analytics for Investigative Analysis
- Situation Surveillance and In-field Criminal Investigative Analytics
- Financial Visual Analytics
- Study of Terrorism and Responses to Terrorism (START Center) Visual Analytics
- Foreign Animal and Zoonotic Disease Visual Analytics
- GeoJunction: Collaborative Visual-computational Information Foraging and Contextualization to Support Situation Awareness
- Video Surveillance Visual Analytics
- Visual Analytics for the DHS Centers of Excellence

**Investigative Analysis and Decision Making Environments**

- Integrate Automated Analytical Reasoning into JIGSAW — a visual index of document collection
- Introduce Sustainable Visual Analytics into Command Center Environments

## BENEFITS OF THE VISIONARY PARTNER PROGRAM

**Visionary Partners** - Contribute $50,000 per year as an unrestricted gift. Gift funds may be designated for general support of the VACCINE Center and to assist with all areas of discovery, learning and engagement, or for work on a specific research initiative.

- Gain early awareness to the latest visualization research. Through meetings and visits, companies get a preview of the latest research findings from across the VACCINE consortium. Companies that can access this information before it's published have a front-row seat to early licensing opportunities.
- Recruit VACCINE students. Our students are our most valuable assets. They add substantial value to any company as interns or employees.
- Provides a neutral convening point for exploring technology futures. VACCINE is the ideal place to bring like-minded people together in a vendor-neutral setting to explore key challenges and technological matters. VACCINE brings to the table the most highly qualified thought leaders from both academia and industry.
- Opportunity to spend a short sabbatical at a VACCINE partner university collaborating on projects of interest while gaining deep insight into the latest visual analytics research and techniques.
- There will be a global goodwill benefit to your business in its association with the 18 VACCINE partner universities.

## OTHER CORPORATE AFFILIATE OPPORTUNITIES

**Project Partners** — VACCINE is very interested in developing research projects under confidential contractual arrangements. These projects typically specify contractually obligated deliverables, address intellectual property issues, and may involve exchange of personnel between sites, as well as regularly scheduled conference calls and updates in addition to written progress reports. Typical projects begin at $20,000 and range upward.

130

**5) In addition to the 18 schools comprising VACCINE, other academic partners**

- Regenstrief Institute
- Mississippi Valley State University
- Center for Infectious Disease Dynamics, Penn State
- San Diego Supercomputer Center
- West Houston Center for Science and Engineering
- Central Washington University
- SUNY Albany
- Penn State Institute for Cyberscience
- Houston Community College
- California State University Dominguez Hills
- University of Maryland
- CREATE
- NCFPD
- CCICADA
- Rochester Institute Technology
- DIEM

**6) VACCINE International Partnerships:**

Utilizing additional funding by DHS and DRDC, VACCINE and the JIBC hosted invitational workshop, "Visual Analytics for Public Safety Professionals". The workshop took place in Canada on September 20-21, 2010 and focused on learning opportunities for public safety professionals and where visual analytics can bring value to their decision-making and daily lives.

VACCINE joined CCICADA, PNNL and our United Kingdom partners in a collaborative SandPit titled "Who Do You Think You Are" held Dec 1-3 and Dec 8-10 in London, England. The focus of this SandPit was to address the kinds of challenges from multiple communities in the area of identity detection, with the goal being to define a challenge area that crossed domains. Around that challenge, the UK and US would fund collaborative research projects to address the agreed-upon areas of concern. These research projects will involve US and UK teams of researchers, with the US components being funded by the US government and the UK performers being funded by the UK government.

- Simon Fraser University, Canada
- University of British Columbia
- Justice Institute of British Columbia
- Swansea University
- University of British Columbia
- University of Stuttgart, Germany
- University of Konstanz
- University of Groningen, Netherlands
- Linkoping University
- Linnaeus University
- University of Gävle
- Chalmers University, Sweden
- National Institute for Research in Computer Science and Control

# VI.  Technology Transition

Throughout this document, it should have become evident that VACCINE is committed to getting our technologies and programs into the hands of the individuals that need them. Since VACCINE's inception in 2009, DHS has gradually moved towards an aggressive technology transition strategy. It is imperative to show value to the end-users via getting VACCINE tools deployed.   As mentioned earlier, the role of developing detailed operational requirements is critical for the successful transition of technologies into the various homeland security fields. The VACCINE Center of Excellence recognizes the need for significant end-user-in-the-loop involvement in the entire life cycle development process.  In 2010, VACCINE created a Public Safety Consortium with the goal of connecting homeland security researchers with the first responder community.  Understanding the requirements of our end-users is key to the research and development of technologies that can be transitioned to the broader public safety community.
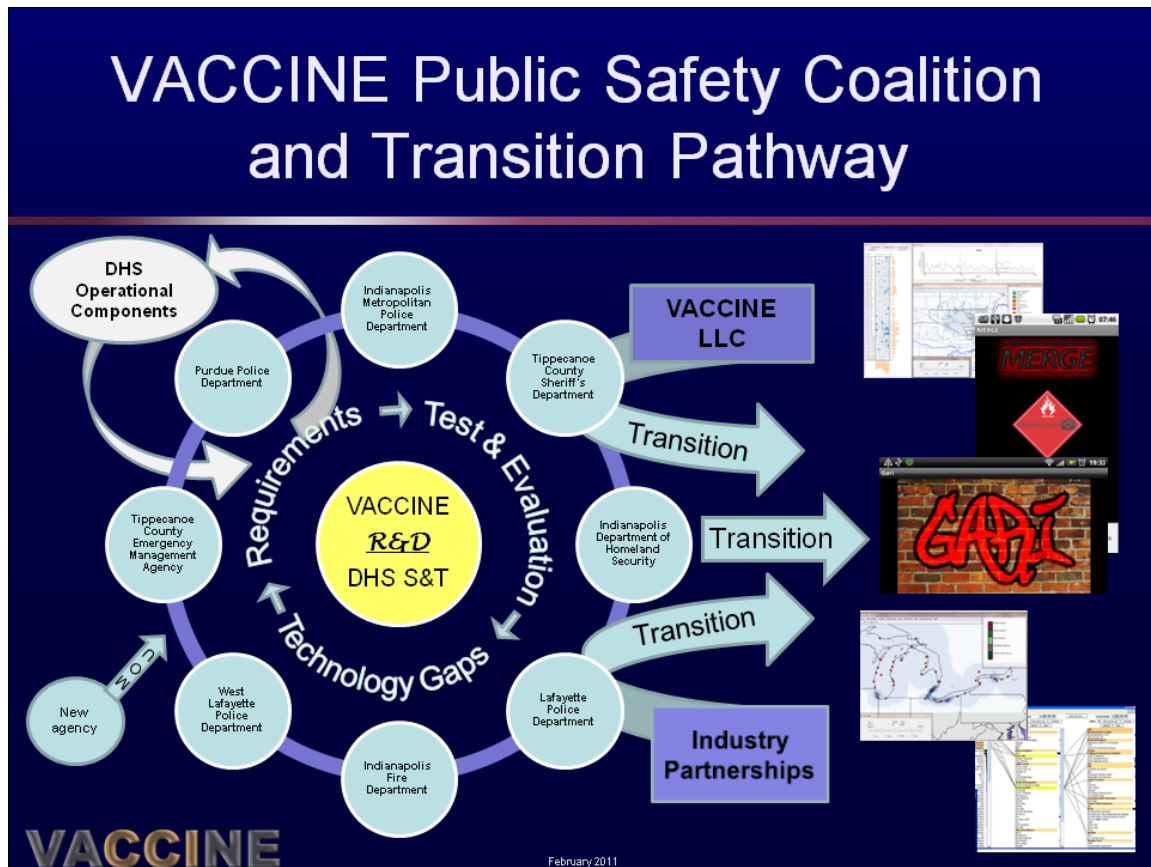


VACCINE Transition Pathway

In 2010, DHS S&T Office of University Programs formed the Research Transition Working Group (RTWG) in order to develop strategies and methods to step up the transition game and get more COE tools, technologies and other products to end users.  The VACCINE Center is represented on the RTWG and has provided significant input into developing the goals and metrics for transition as well as aligning these goals with the recent Quadrennial Homeland Security Review.

Also, as discussed earlier, our industry partnerships and the VACCINE, LLC are critical components of our transition strategy. This is not only true for having potential vendors identified for licensing VACCINE technologies for distribution to broader markets, but also to engage end-users in the education and learning realms. For example, our relationship with Captico, led to the prototype development of the uLearning portal for training and educating the first responder community.

The following list summarizes the VACCINE technologies and where they are in terms of transition and deployment. While they are deployed for evaluation, the feedback we receive has resulted in continuous refinements to meet the needs of the users. In some cases, we are releasing new versions of technologies once a month.

**JigSaw –** VA for Exploring and Understanding Document Collections (Georgia Tech)
- Deployed for Test & Evaluation Feedback Dec - 2010
- West Lafayette PD, Lafayette PD

**PROTECT** - Port Resilience Operational / Tactical Enforcement to Combat Terrorism
- Collaboration with CREATE
- Prototype July 26, 2011

**COE Explorer –** Exploring the Centers of Excellence (Purdue, SFU)
- Collaboration with CCICADA

**CrimeViz -** Sensemaking about criminal activity in space and time (Penn St)
- Testing in progress with Harrisburg PD
- Prototype deployment planned for Q2 2011 Harrisburg PD, JIBC

**ISIS -** Infovis System for Investigating Intrusions (Stanford)
- Awaiting deployment to US CERT

**VALET  -** Visual Analytics Law Enforcement Toolkit
- Deployed for Test & Evaluation Feedback  - 2010
- Lafayette PD, W. Lafayette PD, Purdue PD, Tippecanoe County Sheriff

**GARI -** Gang Graffiti Automatic Recognition and Interpretation
- Initial Prototype Deployment Scheduled  May 2011
- Indianapolis Gang Task Force, Lafayette PD, Purdue PD

**CGSaRVA -** Coast Guard Search and Rescue Visual Analytics
- Deployed  for Operational Use – USCG LANT - 2010

**MERGE -** Mobile Emergency Response Guide
- Initial Prototype Deployment Scheduled May 2011
- Indianapolis Fire, Lafayette Fire

**Ulearning (Captico)-** Training portal for S&T and COE Technologies
- Deployed for Test & Evaluation Feedback - March 2011