"The purpose of computing is insight, not numbers."
—Richard Hamming (1915–1998)

1

Grand Challenges

Our country faces profound challenges that must be addressed in order to ensure our continued freedom and security. As the September 11, 2001, attacks on the Pentagon and World Trade Center illustrate, threats to the United States are present within our borders. On that day, after at least 20 months of planning and preparation, 19 terrorists hijacked four airliners at three different airports in a coordinated attack. The hijackers crashed two planes into the World Trade Center's twin towers and one plane into the Pentagon. The fourth plane, intended to attack another US landmark, crashed in a field in Pennsylvania. These attacks claimed the lives of 2973 people.

As the attack unfolded, government agencies and emergency response personnel "...struggled, under difficult circumstances, to improvise a homeland defense against an unprecedented challenge they had never before encountered and had never trained to meet." They were forced to assess complex and dynamic situations and make decisions rapidly under extreme pressure, often without access to critical information that would help them save additional lives [National Commission on Terrorist Attacks, 2004].

Protecting Our Homeland

Providing for the security of the US and its citizens constitutes a grand challenge for our country. Future attacks could result in major loss of life and shake the nation's confidence in its fundamental security. The US Department of Homeland Security (DHS) was established in 2003 to secure the American homeland and protect the American people. It responds to the nation's security objectives of preventing terrorist attacks within the US, reducing America's vulnerability to terrorism, and minimizing the damage and recovering from attacks that do occur [Department of Homeland Security, 2004].

As described in their Strategic Plan, the mission of DHS is to:

... Lead the unified national effort to secure America. We will prevent and deter terrorist attacks and protect against and respond to threats and hazards to the nation. We will ensure safe and secure borders, welcome lawful immigrants and visitors, and promote the free-flow of commerce.

DHS has identified strategic goals and objectives to accomplish this important mission. Meeting these objectives requires the concerted efforts of professionals in fields as diverse as border and transportation security, intelligence analysis, law enforcement, and emergency preparedness and response (including the firefighting and medical professions). These people are challenged daily with making sense of vast amounts of conflicting and ever-changing information to identify and prevent emerging threats, protect our borders, and respond in the event of an attack or other disaster.

To understand the complexity of this challenge, consider two imagined scenarios.

Scenario 1: Sarin Gas Attack

An organized international terrorist group develops a plot to conduct a Sarin gas attack in the US.

According to the Centers for Disease Control, "Sarin is a human-made chemical warfare agent classified as a nerve agent. Nerve agents are the most toxic and rapidly acting of the known chemical warfare agents." Pure Sarin gas is odorless, so people may not be aware that they have been exposed. Sarin gas attacks the respiratory system and can be fatal [http://www.bt.cdc.gov/agent/sarin/basics/facts.asp].

The terrorist organization sends a planning team to the US to identify the target for this attack. After examining several options, they select three large commercial office buildings in a major metropolitan area. They develop a plan to release the Sarin gas into the ventilation systems of the building. The target buildings are chosen based on their size and location, as well as the relative ease of access to their ventilation systems. The terrorists set up a cell in the target city, and three obtain jobs in building maintenance for the office buildings.

In the meantime, an overseas terrorist cell acquires the Sarin nerve agent, packs it carefully to conceal its true contents, and ships it into the US via a sea-borne cargo container. It passes through Customs at a port city and makes its way to the US-based terrorist cell via rental truck to a safe house near the target city. The overseas cell also acquires protective safety gear and ships it to the US-based terrorist cell.

The local terrorist cell builds the necessary spray dispersal devices, acquiring materials from local stores as needed. On the day of the attack, the terrorists working at the target buildings put on protective equipment and deploy their dispersal devices in the ventilation system, following a strict time schedule to conduct a coordinated attack. Then, as the attackers flee the area, the Sarin gas permeates the air in the three office buildings and escapes though rooftop ventilation systems to the outside air as well.

Within a couple of minutes, people in the three buildings begin to experience the symptoms of exposure. Calls to 911 are placed, and emergency personnel rush to the scene. Initial emergency crews on the scene are unlikely to realize that there has been a terrorist attack and that a nerve agent has been released throughout the buildings, leading them to rush into the building without any protective equipment and suffer exposure themselves.

It may take many minutes, and significant loss of life, before emergency response crews observe enough clues to piece together the true nature of the emergency and identify the agent as Sarin gas. Then they must quickly determine how best to rescue and treat survivors. In addition, because Sarin has been released into the air outside the buildings, emergency managers and city officials must communicate with the public to make clear to them to shelter in place to avoid exposure. Those officials and managers will need to also clearly state the potential for exposure and harm to avoid panic among the public.

Scenario 2: Biological Attack

An international terrorist organization conducts an operation to produce an outbreak of pneumonic plague. Plague is caused by a bacterium, and death will ensue if the infection is not treated rapidly with antibiotics. It can readily spread among humans.

The terrorist organization identifies a major metropolitan area to target for the release of the plague. They select three high-traffic areas as targets: in the bathrooms of the city's major airport, in the bathrooms of its professional sports arena, and in the city's train station.

The plague bacteria are transported into the country via truck at a major border crossing and are driven to the target city. Over the course of one day, local members of the terrorist organization release the plague canisters in the three target locations, timing each release to coincide with the heaviest traffic at each of the locations. The release goes undetected.

Travelers infected in the airport fly to their destinations, both across the US and to other countries. Likewise, tourists in town for the sporting event return to their homes in other locations.

No sign of illness appears until about 36 hours after the initial attack, when infected people rapidly experience severe respiratory illness and die if not properly treated. The rapid increase in unexplained illnesses in the target city triggers alarms within the city's public health system, and public health officials become aware for the first time that an abnormal pattern is emerging. Investigation is required to determine that an abnormal event has occurred. Additional investigation allows the experts to identify that the illness is plague and eventually to tie the cases back to common sources. Multiple government agencies participate in investigations to identify the illness and address the public health needs, as well as to identify the plot behind the outbreak and apprehend those responsible.

While this investigation proceeds, the public in this city are growing worried, and they must be given instructions on the basis of preliminary information to help protect them from illness. By the time the cause is identified, infected individuals who have traveled across the country, and to other countries, are also falling ill and dying, but the reason for the illness will not initially be clear.

Because plague spreads easily among humans, the disease will have the potential to spread far beyond the originally infected population before it is finally identified. Plague has symptoms in common with many milder respiratory illnesses. Part of the major emergency response challenge will be to discriminate the true plague cases from worried members of the public with these milder illnesses. In addition, providing clear and specific information to the public about how to protect themselves from infection will be a high priority.

Grand Challenge: Enabling Profound Insights

The *analysis* of overwhelming amounts of disparate, conflicting, and dynamic information is central to identifying and preventing emerging threats, protecting our borders, and responding in the event of an attack or other disaster. This analysis process *requires human judgment* to make the best possible evaluation of incomplete, inconsistent, and potentially deceptive information in the face of rapidly changing situations to both detect the expected and discover the unexpected. People must collaborate across organizations, agencies, and jurisdictions in a way that allows them to share information when appropriate while adhering to privacy laws and policies.

Employing today's best practices in these areas and using the best possible training, technologies, and processes will still not position the country to address the growing challenges to our security. The volume of data is staggering. Although massive amounts of information are available from multiple sources, the relevant information content exists in a few nuggets. New methods are required that will allow the analyst to examine this massive, multi-dimensional, multi-source, time-varying information stream to make decisions in a time-critical manner. By providing the experts with the capability to truly understand this information, we enable them to make better decisions that prevent and prepare for attacks, improve emergency response, and save lives.

Three major areas deserve particular focus: analyzing terrorist threats, safeguarding borders and ports, and preparing for and responding to emergencies. Strong analytic capabilities are required to accomplish all three of these tasks. Threat analytics is the initial priority, given its importance in the prevention of attacks.

Analyzing Terrorist Threats

As these scenarios illustrate, terrorists take great care to plan and orchestrate their attacks. An urgent goal is to stop terrorist attacks before they occur, which requires that analysts uncover the subtle clues that can identify specifics about the attack before it occurs. To prevent an attack, analysts must piece together details about who is going to conduct the attack, what the attackers plan to do, and where and when the attack will occur.

Observations, such as reports of strangers scouting out the ventilation systems of large buildings, play a critical role in helping to tip off authorities that something is amiss. Analysts need the ability to combine information buried in disparate data, including immigration records; patterns of travel; telephone calls; and names, affiliations, and locations of suspected terrorists, to allow them to piece together the information necessary to spot an emerging attack before it can be executed. In addition, information relating to potential threat agents, the materials and facilities used to make them, and their possible deployment must be considered.

Analysts evaluate potential vulnerabilities that exist in our country and identify and monitor potential threats that may arise. They are also alert for potential anomalies that may indicate that something unexpected is occurring.

Today, analysts have a select number of software programs available to help them organize their information, gain an overview of it, explore it, and examine high-level trends. Although these tools help the analyst, they are only scratching the surface in terms of meeting true analytical needs. Current technologies cannot address the needs for handling the massive, messy, and ever-changing volumes of information and the diversity of types of information. Furthermore, current analytical tools provide basic capabilities, such as query and search, but provide very little in the way of support for the complex tasks of the analysis, synthesis, and discovery process. Very few current tools address the need to communicate analytical results and products to their audiences.

Research is needed to create software that supports the most complex and time-consuming portions of the analytical process, so that analysts can respond to increasingly more complex questions.

Safeguarding Borders and Ports

Safeguarding borders is a complex task. Border guards and customs agents must prevent the illegal entry of goods or people into the country, while allowing them to ensure the free flow of legal commerce.

Daily, thousands of decisions must be made at each point of entry to discriminate between normal, legal activity and potential illegal activity. Tools and processes have been developed to help border and customs agents make these decisions. However, because of the volume of potentially relevant information and the speed with which decisions must be made, we need rapid advancement in the software systems used to support this decision-making process. Individuals in the field need reach-back support to allow them to get more in-depth scientific assessment when necessary so that they can better discriminate illegal from legitimate cargo. At a national level, it is also important to be able to combine information about incidents at ports and borders with other available information to allow analysts to identify patterns that could indicate emerging terrorist threats.

Preparing for and Responding to Emergencies

Even with the greatest of vigilance, attacks can still occur. Emergency preparedness and response are critical to ensuring that, in the event of an attack or any other national disaster, loss of life and property is minimized.

Thorough preparation and response plans have been put in place for a large number of potential national emergencies, and emergency response training and practice drills are regularly held to hone our nation's capabilities. However, terrorist attacks often capitalize on the element of surprise to maximize their effectiveness. As the above scenarios illustrate, in some cases it is not apparent what kind of attack has taken place, or even (in the initial stages) that an attack has occurred at all. This kind of confusion delays the experts' ability to respond effectively. Coordinated attacks may be timed so that a secondary attack occurs once emergency response personnel have arrived, in order to further increase the impact of the attack.

Even with the well-developed emergency plans and procedures that exist, the above scenarios illustrate several areas in which understanding and sharing of information can greatly enhance our ability to respond and greatly reduce the impact of an attack. In the first scenario, for example, real-time analytical monitoring that alerts first responders to the anomalous frequency and location of 911 calls can raise their awareness of the unusual situation in advance. Emergency command centers need analytical capabilities to help them get to the root of emergency situations, rapidly identify the appropriate response, and coordinate ever-shifting facts during a chaotic and pressure-filled time.

Communication is also critical in an emergency situation. First responders must coordinate and communicate clearly across multiple jurisdictions. Sharing of information must be done carefully to protect both privacy and national security. As the scenarios show, communication with the general population is also a requirement. Meeting these communication requirements is time-consuming and challenging, because each audience requires a different type of communication to answer the questions that are most pertinent to them. New software is needed to help streamline the process of customizing communications for different audiences so that people can make better-informed decisions in emergencies.

The Scalability Challenge

Data are growing at an incredible rate. Lyman and Varian [2003] estimate that in 2002 alone, the world produced 5 exabytes ($5x10^{18}$ bytes) of new stored information in the form of paper, film, and electronic media. Another 18 exabytes of streaming information was produced in 2002. Their study estimates that storage of new information is growing at a rate of more than 30% per year.

Analysts, emergency response teams, and border protection personnel have massive amounts of information available to them from multiple sources, but the important information may be hidden in a few nuggets. We must create new methods to allow the analyst to examine this massive, multi-dimensional, multi-source, time-varying information stream to make effective decisions in time-critical situations.

Data Characteristics

Consider some examples of data types that contribute to this information overload.

Textual data. Massive textual data can come from documents, speeches, news, e-mails, or web pages. These data are ever increasing in volume. Our target is to be able to support analysis of data volumes growing at a rate of one billion new structured messages or transactions per hour, and one million new unstructured messages or documents per hour.

Databases. Many corporate and government entities have constructed huge databases containing a wealth of information. These databases are both diverse and distributed. In addition, individuals and workgroups may have their own local databases that augment these large databases. New algorithms are required to permit efficient discovery of previously unknown patterns in these disparate databases.

Image data. Consider the data collected by satellites that image the earth. Commercial satellites can create images at 1-meter resolution and collectively create an image of the planet's land surface in a very short time. New methods are needed to permit efficient understanding of image data, especially in the context of other types of data mentioned here.

Sensor data. The revolution in miniaturization for computer systems has resulted in the production of many types of sensors. The sensors can collect data about their environment (location, proximity, temperature, light, radiation, etc.), can analyze these data, and can communicate among themselves. Collections of sensors can produce very large streaming sets of data. Methods are needed for analyzing sensor data to efficiently incorporate the data into computerized models to support border protection and emergency response.

Video data. Video is often used to enhance the effectiveness of high-risk security operations. Video recording and content analysis are being used in concert as a powerful tool for improving business processes and customer service. New techniques must be developed to integrate these capabilities for analyzing streaming video data into the analyst's toolkit.

The data present challenges not only because of their diversity, volume, and dynamic nature but also because the data are ambiguous, uncertain, and potentially intentionally deceptive. Data of multiple types must often be analyzed in concert to gain insight. Important data needed for correct interpretation may be missing, but this may or may not be apparent to the analyst. We must provide mechanisms that help the analyst visually understand the nature of the data being evaluated.

A grand challenge is to support the analyst in distilling the relevant nuggets of information from widely disparate information streams and create an information space containing relevant information that can be used by the analyst in reaching the most timely and well-informed assessment of the situation. We must provide mechanisms that can visualize the connections among relevant information in the information streams and allow the analyst to gain insight from data.

Scalability

Current technologies cannot support the scale and complexity of the growing analytical challenge. New techniques and underlying scientific foundations are needed to deal with the scale of the problems we are facing in threat analysis, emergency management, and border protection. Issues of scale cut across every aspect of this challenge.

When considering scalability issues, it is important to understand the context of the development of the computer industry as well as natural human skills and limitations. Moore's Law suggests that basic computer technology performance (processor speed and memory density) will double every 18 months. Recently, graphics technology has been improving performance at an even faster rate, doubling every 12 months. This trend has continued for 50 years, and some projections say it will continue for at least another 10 years before fundamental limitations of physics are encountered.

All of this added processing power and memory density has enabled the gathering and processing of vast amounts of data.

However, basic human skills and abilities do not change significantly over time. It is true that technology advances, applied carefully, can enable us to use a higher percentage of natural human abilities, but there are fundamental limits that we are asymptotically approaching. This situation gives rise to the popular notion of *information glut*. That is, we are able to access far more information than we, as humans, can possibly process. The situation also makes scalability issues more difficult to resolve. In addition, analytical challenges often require coping with, sharing, and using information at multiple scales simultaneously. Ultimately, large-scale problems have to be reduced to a scale that humans can comprehend and act on.

Scale brings opportunities as well. For example, increased scale may help reduce uncertainty of an emerging situation. In addition, large data volumes allow analysts to discover more complete information about a situation. As a result, analysts may be able to determine more easily when expected information is missing; sometimes the fact that information is missing offers important clues in the assessment of the situation.

Here, we consider five of the major scale issues that must be addressed: information scalability, visual scalability, display scalability, human scalability, and software scalability.

Information Scalability

Information scalability implies the capability to extract relevant information from massive data streams. Methods of information scalability include methods to filter and reduce the amount of data, techniques to represent the data in a multi-resolution manner, and methods to abstract the data sets.

A second form of information scalability has to do with the rate of change of the information. Most existing techniques do not handle dynamic change, but a few do.

Finally, information presentations must be scaled or adapted to the audience. For example, an analyst's presentation to other analysts will contain far more detail than the summary analysis presented to the President. Current techniques require that this be done manually in an ad hoc fashion.

Relevant information may appear at a variety of scales, and the user must be able to change between scales in a way that is easy to understand and track. We must be able to understand the cross-scale interactions. We must be able to handle a wide range of dynamic change, and we must develop systems that semi-automatically scale or adapt information presentations to match a target audience.

Visual Scalability

Visual scalability is the capability of visualization representation and visualization tools to effectively display massive data sets, in terms of either the number or the dimension of individual data elements [Eick & Karr, 2002]. Factors affecting visual scalability include the quality of visual displays, the visual metaphors used in the display of information, the techniques used to interact with the visual representations, and the perception capabilities of the human cognitive system.

Most published techniques in the field of information visualization handle data sets with hundreds to thousands of elements. Some techniques can scale to handle tens of thousands of elements, and a very few can handle hundreds of thousands up to one million elements. The InfoVis 2003 Contest focused on the problem of visualizing and comparing large hierarchies. The winning technique was TreeJuxtaposer [Munzner et al., 2003], which could handle two trees of about 100,000 elements and one tree up to 500,000 elements.

However, as described previously, some extreme situations may demand the processing of tens of millions of new documents per day, with a total database size of tens of billions of documents. It is reported that at least one existing database has 120 billion documents. It seems likely that these database sizes will increase over time. Clearly the current state of the art is far from able to visually represent today's data collections, and the need will continue to grow. New techniques are needed to bridge this gap.

Display Scalability

Most published visualization techniques are designed for one size display, generally a desktop display (typically 1280x1024 pixels). We need to develop techniques that scale to a variety of display form factors to take advantage of whatever capabilities are available to support analysis and collaboration. Tools should be able to make effective use of everything from a wall-sized display in an emergency response situation room to a PDA or phone-sized display in the hands of a first responder in the field. One major challenge is to develop interaction techniques that are display scale-independent. That is, consistent visualization and interaction techniques should be used regardless of display size. Studies need to be done to determine how to display information effectively, particularly on small displays.

Human Scalability

While human skills and abilities do not scale (i.e., they are relatively fixed), the number of humans involved in analytical problem-solving, border protection, and emergency preparedness and response activities does scale. Most published techniques for supporting analysis are targeted for a single user at a time. We must develop techniques that gracefully scale from a single user to a collaborative (multi-user) environment. Much of the relevant collaboration research is focused on small groups of collaborators (two or three people). In the scenarios we envision, users may be collaborating from within the same team in an organization, at different levels of an organization, or even in different organizations. Each of these cases has its own set of problems that must be solved. One scenario might involve a number of first responders, several regional emergency management centers, and a national emergency management center—that is, dozens of users collaborating through the use of shared analytical tools and focusing on different levels of information accessible by everyone involved.

Software Scalability

Software scalability is the capability of a software system to interactively manipulate large data sets. Software scalability includes the generation of new algorithms that scale to the ever-increasing information sets that we generate today. We wish to avoid the hidden costs that arise when we build and maintain monolithic, non-interacting, non-scalable software.

Other Scalability Issues

Cutting across many scalability issues are concerns with privacy and security, particularly when scaling to multi-user environments. Data privacy and security laws and policies must be adhered to rigorously, which means that software must address challenges such as protecting information from inappropriate access, down to the data item and individual user level. Privacy and security are discussed in more detail in Chapter 6.

Scalability issues also arise in dealing with geographically dispersed teams speaking different languages or using different terminology within the same language, and working across teams of people with differing expertise.

The Need for Visual Analytics

To meet the analytical needs described in this chapter, the scientific community must dramatically accelerate research and development (R&D) efforts to develop fundamentally new solutions. These solutions must enable analysts to focus their full cognitive and perceptual capabilities on their analytical processes, while allowing them to apply advanced computational capabilities to augment their discovery process. R&D in the field of visual analytics helps address these challenges.

Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces. People use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data; detect the expected and discover the unexpected; provide timely, defensible, and understandable assessments; and communicate assessment effectively for action.

Visual analytics integrates new computational and theory-based tools with innovative interactive techniques and visual representations to enable human-information discourse. The design of the tools and techniques is based on cognitive, design, and perceptual principles.

The Research and Development Agenda for Visual Analytics

DHS chartered the National Visualization and Analytics Center™ (NVAC™) in 2004 with the goal of helping to counter future terrorist attacks in the US and around the globe. NVAC is a national resource that provides strategic direction and coordination of activities to discover, develop, and implement innovative visual information analysis methods. A major objective for NVAC is to define a 5-year R&D agenda for visual analytics to address the most pressing needs for R&D to facilitate advanced analytical insight. In spring 2004, NVAC formed a panel of distinguished researchers from academia, industry, and the national laboratory system, as well as select government experts. Through a series of workshops and collaborative efforts, the team established the plan for action summarized in this book.

This agenda builds upon and extends recent government publications, most notably two reports by the National Academy of Sciences. *Making the Nation Safer* [Alberts & Wulf, 2002] describes how science and technology can be used to protect the nation against terrorism. *Information Technology for Counterterrorism* [Hennessy et al., 2003] expands upon the work of *Making the Nation Safer*, focusing specifically on the opportunities for information technology to help counter and respond to terrorist attacks.

Although the agenda described herein is focused specifically on meeting homeland security challenges, the new capabilities created will have an impact on a wide variety of fields ranging from business to scientific research, in which understanding complex and dynamic information is important.

The R&D agenda for visual analytics will require the assembly of a multidisciplinary team to address a set of interrelated research areas illustrated in Figure 1.1.

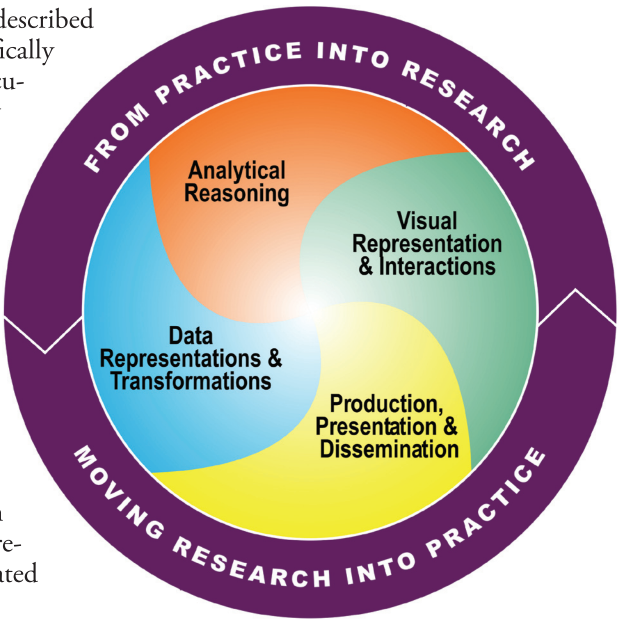

Figure 1.1*. The R&D agenda for visual analytics addresses needs for innovation in interrelated research areas.*

Analytical Reasoning (Chapter 2)

Analytical reasoning techniques are the method by which users obtain deep insights that directly support situation assessment, planning, and decision making. Visual analytics must facilitate high-quality human judgment with a limited investment of the analysts' time. Visual analytics tools must enable diverse analytical tasks such as

- Understanding past and present situations quickly, as well as the trends and events that have produced current conditions
- Identifying possible alternative futures and their warning signs
- Monitoring current events for emergence of warning signs as well as unexpected events
- Determining indicators of the intent of an action or an individual
- Supporting the decision maker in times of crisis.

These tasks will be conducted through a combination of individual and collaborative analysis, often under extreme time pressure. Visual analytics must enable hypothesis-based and scenario-based analytical techniques, providing support for the analyst to reason based on the available evidence. Visual analytics must help the analyst discover the unexpected, whether by detecting unexpected relationships or showing that expected relationships are missing. Visual analytics must also focus on capturing the discoveries and the results of the analytic process and on making them available to others as a basis for future understanding.

Visual Representations and Interaction Techniques (Chapter 3)

Visual representations and interaction techniques take advantage of the human eye's broad bandwidth pathway into the mind to allow users to see, explore, and understand large amounts of information at once. Information visualization research has focused on the creation of approaches for conveying abstract information in intuitive ways. Visual analytics must build upon this research base to create visual representations that instantly convey the important content of information, within context. This visual representation is essential to the analytical reasoning process.

There is no single visual metaphor that can meet all analytical needs. A suite of visual metaphors and associated visual approaches is necessary to provide users with multiple complementary views of their information. Analysts must have the capability to tailor these visual tools to fit their task and their individual analytical style. Approaches are needed to visually represent dynamic data of all types. These representations must be able to convey changing conditions and situational assessments as events transpire, analytical understanding evolves, and requirements change.

Visual representations alone cannot satisfy analytical needs. Interaction techniques are required to support the dialogue between the analyst and the data. While basic interactions such as search techniques are common in software today, more sophisticated interactions are also needed to support the analytical reasoning process. A strong foundation must be developed for interaction science to enable researchers to develop the best interaction techniques to support any given task, timescale, and interaction environment.

Data Representations and Transformations (Chapter 4)

Data are at the heart of the analytical challenge. Analytically important data are buried in vast streams of all types. These data, in their raw form, are rarely appropriate for direct analysis. The key challenge is to create data representations and transformations that convert all types of conflicting and dynamic data into forms that facilitate analytical understanding. These representations must support varying levels of abstraction to facilitate the analysis of massive and dynamic data collections, at multiple scales, and within multiple contexts. Data representations must represent data context where known, but they also must appropriately represent information that is lacking context, incomplete, and uncertain. Today, data are generally analyzed within a collection of a similar data type. Visual analytics must bring all relevant information into a single consistent analytical context, regardless of the form in which the information began, to support analysis and discovery.

Production, Presentation, and Dissemination (Chapter 5)

To have impact, the results of an analysis must be communicated accurately to others. Production of analytical outputs, assembly of presentations appropriate for different audiences, and dissemination of analytical results are essential steps in the analytical process. Results must be communicated to numerous audiences, from policymakers to the general population, in unambiguous and meaningful ways. New approaches are needed to simplify the production of analytical assessments, so that presentations of results, relevant facts and evidence, and associated uncertainties can be assembled rapidly to fit the needs of different audiences. These results must be communicated within the context of the receiver and at the appropriate level of detail to meet their needs.

Moving Research into Practice (Chapter 6)

Supporting these research areas are four important areas that must be addressed to accelerate the often tortuous path from research into practice.

Evaluation. Visual analytics must address methods for evaluating whether or not a particular tool or technique is truly having a positive impact. Usability evaluation, while valuable, is insufficient. Visual analytics must develop meaningful and effective techniques to evaluate the actual value any specific visual analytic technique may provide. Sterile test conditions can often produce different evaluation results than would be observed in actual analytical practice. We must develop evaluation approaches and test data sets that give the best possible indicators of true value in an analyst's hands.

Privacy and Security. Analysts work with information drawn from multiple sources, each of which has associated security and privacy constraints. Laws exist that govern how information may be used and combined, and those laws must underpin visual analytics approaches. Tools must proactively adopt and support approaches such as data anonymization, data minimization, audit trails, and access controls to both protect privacy and ensure information security.

System Interoperability. System interoperability issues are central to the successful research, evaluation, and ultimate deployment of visual analytics tools. Visual analytics tools will be developed by disparate groups, yet they must successfully complement one another in a single seamless analytical environment. Interoperability must be considered at every step in the R&D process, from sharing of research-level code to evaluation of tool suites developed by multiple teams. Visual analytics tools will be deployed in diverse application environments, and they must be engineered to permit flexible implementations that can fit within a variety of application architectures.

Technology Insertion. Technology insertion is the successful transition of a technology into widespread analytical use. This technology transfer process must be planned explicitly. We must understand how new visual analytics techniques fit into the user's overall analytic software environment as well as how they fit into the user's processes. We must define the training paths that help users understand how best to use a new technology and integrate it into their practices.

Fully addressing the research needs described in this agenda will require the concerted efforts of multidisciplinary teams of experts from throughout academia, industry, government, and the national laboratory system. It is imperative to expand the core group of researchers who are addressing these problems. To expand this talent base will require establishing formal educational opportunities for researchers, including educational outreach programs and development of new curricula. Through a set of coordinated research efforts, partnerships, and educational efforts, we can mobilize the research community to address these needs now and in the years to come. Chapter 7 describes the set of initiatives necessary to position the field of visual analytics for enduring success and issues the call to action to accomplish this agenda.

References

Alberts B and WA Wulf. 2002. *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism*. National Academy of Sciences, Washington, D.C.

Department of Homeland Security. 2004. *Securing our Homeland: U.S. Department of Homeland Security Strategic Plan*. Available at http://www.dhs.gov/interweb/assetlibrary/DHS_StratPlan_FINAL_spread.pdf.

Eick S and A Karr. 2002. "Visual Scalability." *Journal of Computational and Graphical Statistics* 11(1):22-43.

Hennessy JL, DA Patterson, and HS Lin, eds. 2003. *Information Technology for Counterterrorism: Immediate Actions and Future Possibilities*. National Academy of Sciences, Washington, D.C.

Lyman P and HR Varian. 2003. "How Much Information?" Available at http://www.sims.berkeley.edu/how-much-info.

Munzner T, F Guimbretiere, S Tasiran, L Zhang, and Y Zhou. 2003. "TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context with Guaranteed Visibility." *ACM Transactions on Graphics*: *Special Issue Proceedings of ACM SIGGRAPH 2003* 22(3):453-462.

National Commission on Terrorist Attacks. 2004. *The 9/11 Commission Report.* W.W. Norton and Company, Ltd., New York.